



451 CTAATCAGAAAGTTACGAGCTGTTATGTATATGACATCCACAAATGC 500  
167 aaenglyilemetgluglyleuglnuproleuasnaaleargvalthr 184  
501 CAATGGTATTATGAAAGCTTGGACACCTTGATGCTATCAGAGTTACAC 550  
184 lnglualevaltrpyrtyrseraspasnaalproilseasnaoproasp 200  
551 AAGAGCGGATGCTACTATTTGATATATGCTCTTAATTTCTAATCCAGAT 600  
201 gluserphelysarvglusergluserasnaulevalserthrsergln 217  
601 GAAAGTTTTAAAGGAGTCGAAAGTAACCTGGTATGACTTCTCAATT 650  
217 userleumetarglinalaleuylsuglnleuileaspproasnleuallat 234  
651 ATCTTTGATCCGTCACAGCTTGAAGCACTGATTTGATCCGAAATTTGGCA 700  
234 hrlysmetprolysglnvalproaspaspheglnleuserillepheglu 250  
701 CTAATATGCCAAACAAAGTCCGATGATTTTCAGCTAAGTATTTTGGAG 750  
251 sergluasplyslyaspplystyrasnlysglytyrglnasnleuense 267  
751 TCTGAGGACAAAGGAGATTAATATTAATAAGATACCAAAATCTTTTGGAG 800  
267 rgllyglleuvalprothrlysprothrprothrprogllyaspproketp 284  
801 TGGTGGTTTGGTCTTACTTAACCACTCCAGAGACCAACCAATGC 850  
284 robproasnlnproglnthrthrservalleuilearglystyrallale 900  
851 CTCCAATCAACTCAAAACGACTTCACTTAATTAAGAAATGATATGCTATA 900  
301 glyasptryserlyseuleuuglyalathrleuclnleuthrglyas 317  
901 GGATATTACTCTAAATTCCTTGAAGGTGCAACATYACAGTTGACGGGGA 950  
317 pahnvalasnserpheglnalargvalpheseraserasnaspillegly 334  
951 TAAAGTCAATAGTTTCAACGAGAGCTTTAGCAGTAATGATTTGGAG 1000  
334 luargileuileuSeraspilythrtyrthrleuthrgluLeuasnSer 350  
1001 AAGAAATTTGAATCATCATGATGGAAGCTTACTTTAATGAAATGAAATCT 1050  
351 proalaglyrserillealagluproilethrphelysvalglualagl 367  
1051 CCAGCTGGTTATAGTATCCGAGAGCCAAATCACTTTAAGGTGAAAGCTGG 1100  
367 ylysvalltyrthrilleaspglylsglnilegluasnproasnlysg 384  
1101 CAAAGGTATATCTATTTATGATGGAACACAGATTGAAATCCCAATTAAG 1150  
384 luilevalgluprotyrservalglualatyraasnaspheglnlupe 400  
1151 AGATAGTAGAGCTTACTACTAGTAAGCAATATATGATTTTGAAGATTT 1200  
401 servalleuthrthrlnasnlytyralalyspheytyrtyralalysanly 417  
1201 AGCGTTTAACTACACAAAACATATGCAAAATTTTATATGCAAAAATTA 1250  
417 sasnglyserserglinalvaltyrcyspheasnaalaspdeuylserp 434  
1251 AAATGGAAGTTCACAGGTGTCTATTTCTTTAATGCAATCTAATAATCTC 1300  
434 robproaspsergluaspglyltyrlysthrmethrproasphefthrthr 450  
1301 CACACAGCTCTGAAGATGCTGGGAAACAAATGACTCCAGACTTTACAACA 1350  
451 glygluvallystyrthrhisillealaglyargaspheulehlystyrth 467  
1351 GGAGAGTAAATAACACTCATATTTGACAGTCTGACCTTTTAAATATATAC 1400

467 ryallysproatrgaspthraspproaspthrphleuylshisillelysl 484  
1401 TGTGAACCAAGAGATACCGATCTCGACACTTTCTTAAAAACATATCAAAA 1450  
484 ysvalilleglulysglytyrargglulysglylnalalleglutyrsr 500  
1451 AAGTAATTTGAAAGGTTTACAGGAAAAAAGACACACTATTTGATATAGT 1500  
501 glyleuthrglnthrclnleuargalalaathrglnleuallatletyry 517  
1501 GGTCTAATCTAGACACAATTTGCGTGGCTACTACGTTACCAATATATATA 1550  
517 rpherthraspseralagluLeuaspplysaspplysleuylasptryhis 534  
1551 TTTCACTGATAGTGTGAATTTAGATTAAGATTAACATAAAGACTATACATG 1600  
534 lyphleglyaspmetasnaspserthrleualalvalalylsilleuval 550  
1601 GTTTTGGAGCATGAATGATATGATCTTTAGCAGTGTCTTAAATCTTGTGA 1650  
551 glutyrallaglnaspserasnproglinleuthraspleuaspheph 567  
1651 GAATACGCTTAACATAGTATATCTCCACAGCTAAGTACCTTATTTCTT 1700  
567 eileproasnasnasnlystyglnserleuileglythrglntrphisp 584  
1701 TATTCCGAATTAACATTAATATCAATCTTTATGCAACTGATGCGATC 1750  
584 rogluaspleuvalaspilleleargmetgluasplyslysgluvalille 600  
1751 CAGAGATTTAGTTGATATTTATTCGTATGAGATTAATAAAGAGTTATA 1800  
601 provalthrhisasnleuthrleuarglysthrvalthrlyglleuallagl 617  
1801 CCTGTAACTCATATTTAATTTACATTGAGAAAAACCGTGACTGTTAGCTGG 1850  
617 yaspargthrlyasphehispheglnillegluleuylasnasnlysg 634  
1851 TGACAGAACTAAAGATTTCCATTTTGAATTTGAATTAATAAATATATAGC 1900  
634 lngluLeuLeuSerGlnThrVallysthraspplysthrasnleuglupe 650  
1901 AAGAAATGCTTTCTCAAACTGTTAAACAGATTAACCAAACTCGAATTT 1950  
651 lysaspglylysallathrilleasnleuylshisglyluserleuthrle 667  
1951 AAAGATGCTAAAGCAACCATTTAATTTAAACATGGGAAAGTTTAACACT 2000  
667 uginglyleuuprogluuglytyrserlyrleuvallysgluThraspSerg 684  
2001 TCAAGTTTACCAAGAGTTATTTCTTACTTGTCAAAAGAAACAGATTTCTG 2050  
684 lnglytyrlyvallysvalasnserglnluvalalaaalathrval 700  
2051 AAGCTATTAAGTTAAAGTTAATAGCCAAAGATGAGAAATGCTACAGT 2100  
701 serlysthrcllylethrseraspgluThrleualaphegluasnasnly 717  
2101 TCAAAAACAGGAATTAACATGATGACACACTTCTTTTAAATATATAA 2150  
717 sglnprovalvalprothrlygllyalaspilnlylileasnglytyrleua 734  
2151 AGAGCTGTGTTCTCTACAGAGTGTATCAAAAGATCAAAAGCTTACTAG 2200  
734 laleuilevalillealaglylleserleuuglylletppllyllesthr 750  
2201 CTTTGAATGATTAATCGTGTATCAGTTTGGGATCTGGGAAATTCACAG 2250  
751 lleargtlllearglyshisasp 757  
2251 ATAGGATTAAGAAACATGAC 2271

```

seq_name: /cgnl_7/pdata/2/pna/US09_NEW_COMB.seq:us-09-494-297-3
seq_documentation_block:
; Sequence 3, Application US/09494297
; GENERAL INFORMATION:
; APPLICANT: PODBIELSKI, ANDREAS
; TITLE OF INVENTION: COLLAGEN-BINDING PROTEINS FROM STREPTOCOCCUS PYOGENES
; FILE REFERENCE: P0662850/BAS
; CURRENT APPLICATION NUMBER: US/09/494,297
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2229
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-09-494-297-3

```

```

alignment_scores:
Quality: 1965.50      Length: 737
Ratio: 3.412          Gaps: 13
Percent Similarity: 78.155      Percent Identity: 54.410

```

```

alignment_block:
US-09-494-297-2 x US-09-494-297-3 ..

```

```

Align seg 1/1 to: US-09-494-297-3 from: 1 to: 2229

```

```

20 SerIysaSerIysArg.....PheThrValIhrIeuValIglValAlp 34
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
31 AGCCCTAACACAAACGACGACAAACGATCGATTACTGAAAGTATT 80
34 eIeMeTiIePheAlaIeuValIhrSerMeIValIglValIsthrValp 51
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 TTTGACGTTTCTACTCTGATAGAGATAGAGTGGTTTCTATCAGAGCGT 130
51 heGIleuValIgluSerSerThrProAsnAlaIleAsnProAspSerSer 67
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 TCGGAGCTGAAGAACATCAGTACCAAT.....AGACAAACCTCAATT 174
68 SerGIuIyrArgTrpIyrGIyTrpGIuSerTrpValArgGIyIsthrPro 84
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
175 CAAGATTATCGGTGATGCTATGATCTTATCTAAAGGCTACCCAGA 224
84 rTyIySGIuPheArGValAlaHISAspIeuArGValAsnIeuGIyS 101
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
225 CTATAGCCCGTTAAAGACTTACCTATTAATTTAAAGTAAATTTAGAGGAA 274
101 eArGSerTrpGIuValIyrcysPheAsnIeuLysLysAlaPheProIeu 117
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
275 GTAAGGATTATCAAGCATACTGCTTATTTAAACAAACATTTCCATCC 324
118 GIySerAspSerSerValIySTrpIyIySLSHISAspGIyIleSe 134
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
325 AAGTCAATAGTGTAGATCAATGATATATAAAACCTGAAGGAACCTAA 374
134 rThIySPheGIuSPTyrAlaMetSerProArGIleThrGIyAspGIu 151
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
375 TGAAGAACTTATCAAGATTAGCAGATTAACCAAGATAGACAGCAGAGT 424
151 euAsnGIuLysLeuArGAlaValIeTrpAsnGIyIsthrProGIuAsnAla 167
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
425 TACAAACAAATATATGAGATCTCTATATATGATATCCATTAATTCGT 474
168 AsnGIyIleMetGIuGIyLeuGIuProIeuAsnAlaIleArGValThrGI 184
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475 AATGGGATTAATGAAGGATAGATCTCTAAACGCTATTTTAGTGATCA 524
184 nGIAlaValArTrpIyTrSerAspAsnAlaProIleSerAsnProAspG 201
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
525 AATGCTATTG...TATACTGATTCAGCTCAATTT...AATCGGATG 568

```

```

201 IuSerPheLysArgGIuSerGIuSerAsnIeuValSerThrSerGIuIeu 217
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
569 AAAGTTTTAAACAGAGCTGAGATATGCTATTATATGACACACAGCTTA 618
218 SerLeuMetArgIuAlaIeuLysGIuIeuIleAspProAsnIeuAlaTh 234
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
619 GCCTTAATGCGAAGAAAGCTTAAAGAACTATATGATCAAACTTAGGGTC 668
234 rIyMetCProLysGIuValProAspAspPheGIuIeuSerIlePheGIuS 251
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
669 AAAATATTCGAATTAACCTCCATCAGTTATCGGTAAATGATATTGAT 718
251 eArGIuAspLysGIyAspLysTrpAsnLysGIyTrpGIuAsnIeuLysSer 267
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 CTCATGAT.....AAGCCTTCCAAATCTTTTGAGT 750
268 GIyGIyLeuValProThIySProProThrProGIyAspProPromeItr 284
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 GCTGAGTATGTTCCGGAATCTCCCAAAACGAGAGAGAG.....CC 794
284 OProAsnGIuProGIuIhrThrSerValIeuIleArGIySTyrAlaIleG 301
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
795 TCCCGTAAACCTGAAAAAACATCAGTCATTTATCAGAAATATATCGGAG 844
301 IyAspTrpSerLysLeuIeuGIuGIyAlaThrIeuGIuIeuThrGIyAsp 317
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
845 GTGAC...TCTAACTCTAGAGGAGCAGACCTTAACCTTCTCAATT 891
318 AsnValAsnSerPheGIuAlaArGValPheSerSerAsnAspIleGIy 334
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
892 GAAGAGAGTGGTTTCAAGAAAGAACTTTCAAGTAAATATTTAGGAGA 941
334 uArGIleGIuIeuSerAspGIyTrpTrpIhrLeuThrGIuIeuAsnSer 351
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
942 AACTGTCAATTTACCAATGAGGACTTATCTTAAAGAAACATCATCTC 991
351 rOAlaGIyTrpSerIleAlaGIuProIleThrPheLysValGIuAlaGIy 367
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
992 CAGATGATATTAATAATGCGGAGCGATTAAGTGTAGAGAAATATAA 1041
368 LysValIyTrpIle...IleAspGIySGLIleGIuAsnProAsnLys 383
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1042 AAAGTATTTATTCGCCAAAAAGATGTTCTCAAGTGGAAAAATCCAAACA 1091
383 sGIuIleValGIuProIySerSerValIgluAlaIyTrpAsnAspPheGIu 400
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1092 AGAAGTACAGAGCCATATCTCAGTGAAGCGTATATATGACTTTATGATG 1141
400 heSerValIeuThr...ThrGIuAsnTrpAlaLysPheTrpTrpAlaLys 415
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
1142 AAGAGTACTCTCGGGTTTACTCCATCGGAAAAATCTTATAGCTTACA 1191
416 AsnIySAsnGIySerSerGIuValIyrcysPheAsnAlaAspIle 432
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1192 AATTAAGGTTAAAGTTCACAAAGTTGTCTACTGCTTCAATGCTGATTACA 1241
432 sSerProProAspSerGIuAspGIyGIyLysThrMetThrProAspPhe 449
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1242 CTCACACCTGATCATATGATAGTGGTGAAGCTATAATACAGATACTA 1291
449 hrThr...GIyGIuValIySTyrTrpHisIleAlaGIyArGAspIeuPhe 464
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1292 GTACAGATCAAGAAAGTCAAGTACACACATACGCGAGTACGATCTTGT 1341
465 LysTrpTrpValIySProArGIAspTrpAspProAspThrPheLysSHI 481
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1342 AATATATGCTTAGACCGAGATACAAATCCAGAGACTTCTTAACCA 1391
481 sIleLysLysValIleGIuLysGIyTrpArGIuLysGIuAlaIleG 498
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1392 CATTAATAAAGTAAATGAAAAAGGCTACAGAAAAAAGGTATATGAC... 1437
498 IuTrpSerGIyLeuThrGIuThrGIuIeuArGAlaIleThrGIuIeuAla 514

```



4078 GTCACATAAAGACATACGAACTGCAACT.....GCAGATCAG 4038  
 194 AlaProIleSerAsnProAspGluSerPheIysArgIleuSerGluSerAs 210  
 4037 CCTAAAGTGCATAGTACTGATAGTCAACAGTGTAAAGAACTAGTAGTAA 3988  
 210 nIeuValSerThrSerGlnIleuSerIleuMetArgIleAlaIleuIysGln 227  
 3987 CATGCATTCACACCAAAACGCTACAGCTAATCAATCT..... 3951  
 227 euIleAspProAsnIleuAlaThrIysMetProIysGlnIValProAspAsp 243  
 3950 .....ACTACAAAACCTAGCATGTAAACACAAACATGAT 3918  
 244 PheGlnIleuSerIlePheGluSerGluAspIysGluSplysIyrAsnIly 260  
 3917 AAATCATCACTACATATAGTAAATGAAGAACT.....GATAAA..... 3882  
 260 sGlyTyrgIlnAsnIleuSerGlyGlyIleuValProThrIysProIle 277  
 3881 .....AGTAATTTAACACAGCAAGCAAAAGATGTTTCACTACACCTAAA 3839  
 277 hrProGlyAspProIleProMetProProAsnGlnProGlnThrThrSerVal 293  
 3838 CAACGACTAATTAACCAAGAACTTAAATCGCATGGCAGTGAATACGTGT 3789  
 294 IeuIleArgIysIyrAlaIleGlyAspIyrSerIysIleuIleuGluIylAl 310  
 3788 GCACGCTCCACA..... 3777  
 310 aThrIleuGlnIleuThrGlyAspAsnValAsnSerPheGlnAlaArgVal. 326  
 3776 .....CAAGAACAAATGTTAAATGAT.....AAGGTAC 3749  
 327 .PheSerSerAsnAspIleGlyGluArgIleGluIleuSerAspGlyThr 342  
 3748 ATTTTCAAATATGACATTCGCGATGATAAGCAACATGTTAAATCAGACT 3639  
 343 TyrThrIleuThrGlu.....IleuAsnSerPr 351  
 3698 ACTGCTAAACCTGAATTTTGGCAACTTCAAGTGATGTTTAAATTTAA 3649  
 351 oAlaGlyIyrSerIleAlaIleuProIle.....ThrP 362  
 3648 AGCAAAATTCACATCGATGATCTGTAAAGAGCGCATACATTTACTT 3599  
 362 heIys.....ValGluAlaGlyIysValIyrThrIleIleIleAsp 374  
 3598 TTTAAATTTGGTCAATATTTCCGTCAGATCAGATCAATTAACCTTCACA 3549  
 375 GlyIysGlnIleGluAsnProAsnIysGluIleValGluProIyrSerVa 391  
 3548 ACTCAAAATTTATATATGCCCAAGTAAATATATTT.....GCAAA 3508  
 391 IglAlaIleAsnAspPheGluIleuPheSerValIleuThrThrGlnAsnT 408  
 3507 AGGTATTTATGATAGTACAAACAAACACACACATATATCTTTAGCAACT 3458  
 408 yAlaIysPheIyr.....Tyr 413  
 3457 ATGTAGATCATATATACAAATGTTAGAGTACCTTTGAACAAAGTTGCATTT 3408  
 414 AlaIysAsnIysAsnGlySerSer..GlnValIValIyrCysPheAsnAl 429  
 3407 GCGAAAGGTAAATTCACACACATGATAAACAGCTTATAAATGAGAGT 3358  
 429 aaSPleuIysSerProProAspSerGluAsp.....GlyG 441  
 3357 AACTTATGATGATATACATATAGCAAGAAATCATGTCGATATATGTTA 3308  
 441 IlyIysThrMetThrProAspPheThrThrGlyGluValIysIyrThrHis 457  
 3307 ATAAAAAAGCAACACCGCTTATTTCAGTACAAACATATATTAAACATGAA 3258

458 IleAlaGlyArgAspIleuPheIysIyrThrValIysProArgAspThr.. 473  
 3257 GATTATTCGCGTAATATATGACTGATATGTAATTAACCTAAAAATACATA 3208  
 474 .AspProAspThrPheIleuIysHisIleIysIysValIleGluIysGlyT 490  
 3207 TACTAAACAAACGTTGTACTAATTTA.....ACTGAT 3173  
 490 yTrArgIuIysGlyGlnAlaIleGluIyrSerGlyIleuThrGluThrGln 506  
 3172 ATAAATTTAATCCAAATGCAAAAAACCTCAAA..... 3141  
 507 IeuArgAlaIleThrGlnIleuAlaIleIyrThrIyrPheThrAspSerAlaG 523  
 3140 .....ATTACGACTGACAGATTCAAAATCA 3115  
 523 uIleAspIysAspIysIleuIysAspIyrHisGlyPheGlyAspMetAsnA 540  
 3114 ATTT..... 3111  
 540 sPserThrIleuAlaValAlaIysIleIleuValGluIyrAlaGlnAspSer 556  
 3111 ..... 3111  
 557 AsnProGlnIleuThrAspIleuAspPheIleProAsnAsnAsnIly 573  
 3110 .....GTGGATAGTTTCACCCCTGATCTCAMA 3082  
 573 sTyrgIlnSerIleuIleGlyThrGlnIyrPheIleProGluIleuValAspI 590  
 3081 ACTTAAAGATGTT.....ACTGATCAATTCGATG 3053  
 590 IeIleArgMetGluAspIysIysGluValIleProValThrHisAsnIleu 606  
 3052 TTATTTATAGTAATGATATATAA.....ACAGCTACAGTC 3018  
 607 ThrIleuArgIysThrValThrGlyIleuAlaGlyAspArgThrIysAspPh 623  
 3017 GATTTAATGAAAGGCCAACA..... 2997  
 623 eHisPheGluIleGluIleuIysAsnAsnIysGlnIleuIleuSerGln. 639  
 2996 .....ACGACCAATTAACATACATCATTCACACAG 2966  
 640 .....Thr 640  
 2965 TTGCTATCCAGATATAGTTCACACAGATATGAAAAATGATTACT 2916  
 641 ValIysThrAspIysThrAsnIleuGluPheIysAspGlyIysAlaThrIle 657  
 2915 TTAGACATCGCAAAACCTAAATATAGTTGGTCAAAATGTTATTCMAATGT 2866  
 657 eaSPleuIysHisGlyIleuSerIleuThrIleuGlnGlyIleuProGluIyl 674  
 2865 GAAT.....GCCTCATCACTGCTAATGGGACCAAAAGAAAT 2828  
 674 yrSer.....TyrIleuValIysGluThrAspSerGluIylIys 687  
 2827 ATATATCAGTACTATGATGGAACATACAAATTAAGATGT..AAA 2781  
 688 ValIysValAsnSerGlnIleuAlaIleAsnAlaThrVal.....Se 701  
 2780 CAAGATCCCATGAAAAAGGATTAAGGTGTTATGTCATTTCTTAAGA 2731  
 701 rIysThrGlyIleThrSerAspGluThrIleuAlaPheGluAsnAsnIys 717  
 2730 TAGTACGCTTAAGAAATTAATGATCGTACACACAGATGAAAATGTAAA 2682  
 seq\_name: /cgnl\_7/plodata/2/pna/us08\_NEW\_COMB.seq:us-08-956-171C-472  
 seq\_documentation\_block:  
 ; Sequence 472, Application US/08956171C

GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
Gill H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171C  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoover, Kenley K.  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 610-5790  
FAX: (301) 309-8439  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6806 base pairs  
STRAND: double  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 472:  
US-08-956-171C-472

alignment\_scores:  
Quality: 136.50 Length: 794  
Ratio: 0.403 Gaps: 39  
Percent Similarity: 42.695 Percent Identity: 19.270

alignment\_block:  
US-09-494-297-2 x US-08-956-171C-472/rev ..

Align seg 1/1 to reverse of: US-08-956-171C-472 from: 1 to: 6806

137 PhgGlnAspIyrAlaMetSerProArgIleThrGlyAspGluLeuAsnG1 153  
|||||  
6573 TTTAAATGATTATGACGCT.....CGTATT.....ATGAACTTAAGC 6536  
153 nlySLeuArgAlaValMetYrAsnGlyHisProGlnAsnAlaAsnGlyI 170  
|||  
6535 AAAAGATAAAAACTTA.....GATGGTATGATA 6507  
170 leMetGluGlyLeuGluProLeuAsnAlaIleArgValThrGlnGluAla 186  
|||  
6506 TTGCTGAAGCTTAAACAGCTGTGTGTCTGATTCAGAAAGAAATTA 6457  
187 ValTrpTyrIyrSerAspAsnAlaProIleSerAsnProAspGluSerPh 203  
|||  
6456 TTG.....CAATT 6449

203 eLySArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuM 220  
|||||  
6448 TGAAGGACAAACGAATCTAAATTG...GGTACTTCTGAA..... 6412  
220 eLArgGlnAlaLeuYrSglnLeuIleAspProAsnLeuAlaThrLysMet 236  
6411.....GCTAGAAAGCTGTGTGATTCAGTTGTTGACAGCAAAATTG 6373  
237 ProLYsGlnValProAspAspPheGlnLeuSer..... 247  
6372 CCATTCTATTAGAGAAAAAGCAATGTTCTAAATCACTTGTGA AAAA 6323  
248 .....IlePheGluSerG 252  
6322 AGCGATTAAAGCACAAACAGCAAGGAGCTGCACGTAAAGCTGGAG 6273  
252 LuAspLYsGlyAspLYrTYrAsnLYsGlyTYrGlnAsnLeuSerGly 268  
|||  
6272 ATGCTGCTTCAGGTAAGAAAAACAGCGTAAACACTTGTCTATCTGCT 6223  
269 GlyLeuValProThrLYsProProThrProGlyAspProPheProPr 285  
|||  
6222 AAATTAAACCTGCACAA.....AGTAA 6200  
285 oAsnGlnProGlnThrThrSerValLeuIleArgLYrAlaIleGlyA 302  
|||  
6199 AAACACTGAAAAAATGAAATGATTGATTAGTCGAAGGTGATTCGCGGAG 6150  
302 spTYrSerLYsLeuGluGly.....AlaThrLeuGlnLeu 314  
|||  
6149 GTTCAGCAAAACTTGACGACGACCGCAATTCACAGCGATATTACCAATTA 6100  
315 ThrGlyAspAsnValAsnSerPheGlnAlaArg.....ValPheSe 328  
|||  
6099 CGTGGTAAGTATTAATACAGAAAGCAGCTGTAAGATATTATTAA 6050  
328 rSerAsnAspIle.....Glyc 334  
|||  
6049 AAATGAAGAAATTAATACATTAATCCACACAAATCGGGCAGCGCTTGTA 6000  
334 LuArgIleGluLeuSerAspGlyTYrTYrThr.....LeuThr 346  
|||  
5999 CTGACTTTAAATGAAGATGATTAATATATATCGTGAATATTATAGACT 5950  
347 GluLeuAsnSerProAlaGlyTYr..... 354  
|||  
5949 GATGCTGATCTGATGTGGCATATTCAGAGTGTATTTACATTCTT 5900  
355 .....SerIleAlaG 358  
5899 CTTCAAATATATGAAACCGCTGTCTTCAGCAGGTGCTGATTATTGCTT 5850  
358 LuProIleThrPheLYsValGluAlaGlyLYs..... 368  
5849 TACCTCCACTTTATAAATGGAAGGTAAGGCAAAACAAACCGAGTT 5800  
369 ValTYrThrIleLeAspGlyLYsGlnIleGluAsnProAsnLYsGluI 385  
|||  
5799 GAATACGCTTGACAGC...GAAAGCTTAAATAATTCGAAAAAGAACT 5753  
385 eValGluProTYrSerValGluAlaTYrAsnAspPheGluGluPheSerV 402  
|||  
5752 TGGTAAAGCTTCACGTTACAGCTTAAGAGGTTTGGTGAA..... 5710  
402 alLeu.ThrThrGlnAsnTYrAlaLYs..... 410  
|||  
5709 ..ATGAACCTGGAACAAATTATGGGAAACGACGATGAACCAAGAAACGA 5662  
411 .....PheTYrAlaLYsAsnLYsAsnGlySerSerGlnVal. 423  
|||  
5661 ACTTATATGCTGTACAAAGTTGAAGATGAAGTGGCTCATCTAAACGTGT 5612  
424 .....ValTYrCysPheAsnAlaAspLeuLYsSerProPro 435



52 LyleValGluSerSerThrProAsnAlaIleasnProaspSerSerSer 68  
||| : : : : :  
91 . TTGCAGAAAAAATCTTGGCTTTTCAGAAAGTAATAATTATACCACTAGGC 138  
69 GluTyArgTrpTyrGlyTyrGluSerTyrValArgGlyHisProCysLeu 85  
||| : : : : :  
139 GATTCCGTACACTCGATGATTATGAAGAACA. .... TA 173  
85 rLysGluPheArgValAlaHisAspLeuArgValAsnLeuGluGlySera 102  
||||| : : : : :  
174 CCGGATTTTTCAGTGACTTTAGTAAAGTGGTAATTTTAAGCGGT...C 220  
102 rGSetrYrGlnValTyr.. CysPheAsnLeuLysLysAlaPheProLeu 117  
:||| : : : : :  
221 AATCTTATCTCGCTTATTTTGTATTATATCTACAAA...CTT 261  
118 GlySerAspSerSerValLysLysTrpTyrLysHisAspGlyLeu 134  
: : : : :  
262 AATAGAAATTCACCTGTT..... TCTATGATAAATTTGGC 296  
134 r..... ThrL 136  
297 AATTTCTGGACAAACAACAGAAATCTACTTACCCTTAAATCCAACCA 346  
136 yspheGluAspTyrAlaMetSer.... ProArgIleThr..... 147  
||| : : : : :  
347 AATATCCAAATGAGAAATGACGATATACCTTTAGTTACAAACTATTTCA 396  
148 GlyAspGluLeuAsnGlnLysLeuArgAlaValMetTyrAsnGlyHisPr 164  
||| : : : : :  
397 GGAAATGCAAATAATATGAAATAGGTCGTGTTTTGGCATTTAAATGA 446  
164 OGInAsnAlaAsnGlyIleMetGluGlyLeuGluProLeuAsnAlaIleA 181  
: : : : :  
447 GCATCTCTATCCCTGGCTTAGTCCAAAAAGTTAAGAAAGCAAACTTTTCA 496  
181 rGyalThrGlnGluAlaValTyrTyrTyrSerAspAsnAlaProIleSer 197  
: : : : :  
497 CAATGTCAGTGGAGCT..... AATGATCCTTTTGA 528  
198 AsnProAspGluSerPheLysArgGluSerGluSerAsnLeuValSerTh 214  
529 GCATTTTATGATGATTTTAAAAATGAGCA..... 558  
214 rSerGlnLeuSerLeuMetArgAlaLeuLysGlnLeuIleasPrPro 231  
: : : : :  
559 AGATATATTAACCAAAATCACAGAGAACAAAAAATTAATCAATGCCAA 607  
231 snLeuAlaThrLysMetProLysGlnValrProAspAspPheGlnLeuSer 247  
: : : : :  
608 AT..... GAAGAGCCGAATTTCTGCGACAGAAAAAGAAATGCTTTTA 648  
248 IlePheGluSerGluAspLysGlyAspLysTyrAsnLysGlyTyrGlnAs 264  
||| : : : : :  
649 AAAGCGAATCAATAAAAAATTAAGAAATTAACACAAATCTTGATTA 698  
264 nLeu..... LeuSerGlyGlyLeuV 271  
||| : : : : :  
699 TTTAATTAAGAAATTAAGCGCTTAATCCAAATTAAGTATTAATTA 748  
271 alProThrLysProProThrProGlyAspProProkMetProProAsnGln 287  
:||| : : : : :  
749 TTGGATATTAATTTGCCAAATCCGGT..... 774  
288 ProGlnThrThrSerValLeuIleArgLysTyrAlaIleGlyAspTyrSe 304  
: : : : :  
775 ..... TTATTAAGATTTTAATAATCTTTATATACCTTAATGC 812  
304 rLysLeuLeuGluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsn 321  
||| : : : : :  
113 AAAA... ATTGAAACGACTTATATCAATGAAATTCGCCGAAAAAATTAACA 859

321 erPheGlnAlaArgValPheSerSerAsn..... AspLe 332  
: : : : :  
860 AATATATCTCGTGAACGCCCATTAATAAATGAAGTAATATATATGATGTC 909  
333 GlyGluArgIleGluLeuSerAspGlyThrTyrThrLeuThrGluLeuAs 349  
: : : : :  
910 TAGATATTAAGTATTTGAATGATTTCTGATTAATAATTAATGCGAAAA 959  
349 n..... SerProAlaGlyTyr... SerIleAlaGluProI 360  
||| : : : : :  
960 TTTTGACTTCCACCTTCATTCAGGTTATTAATAAATGCTCACCAAC 1009  
360 lerThrPheLysValGluAlaGlyLysValTyrThrIleIleAspGlyLys 376  
:||| : : : : :  
1010 TTTTGTATTA... CTTCACCAACAAA 1032  
377 GlnIleGluAsnProAsnLysGluIleValGluProTyrSerValGluAl 393  
: : : : :  
1033 GAAAGATGATCTTCAATGCTGTA..... GAGCTAAATAAATCACTACAAA 1076  
393 aTyAsnAspPheGluGluPheSerValLeuThrThrGlnAsnTyrAlaL 410  
:||| : : : : :  
1077 TTTGATGATTTTGATGAC..... 1095  
410 yspheTyrTyrAlaLysAsnLysAsnGlySerSerGlnValValTyrCys 426  
||| : : : : :  
1096 ..... AATTAACCGACCTANTTCCAAAGTATTT... 1122  
427 PheAsnAlaAspLeuLysSerProProAspSerGluAspGlyGlyLysTh 443  
1122 ..... 1122  
443 rmetThrProAspPheThrThrGlyGluValLysTyrThrHisIleAlaG 460  
1122 ..... 1122  
460 LysArgAspLeuPheLysTyrThrValLysProArgAspThrAspProc 476  
1123 ..... GACCTAAGT 1131  
477 ThrPheLeuLysHisIleLysLysValIleGluLysGlyTyrArgGlyL 493  
: : : : :  
1132 GTTTTGCAAAATCAAAATTAAGATTTCTTGAAAAA... TTAACGAAAA 1178  
493 sGlyGlnAlaIleGluTyrSerGlyLeuThrGluThrGlnLeuArgAla 510  
: : : : :  
1179 TAAGCAAACTAGATGATTTATGCTCAAAAAATCCACT..... 1215  
510 lathrGlnLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAspLys 526  
: : : : :  
1216 ..... TTGACACCGCATCAAGAACTGCAATCAAA 1245  
527 AspLysLeuLysAspTyrHisGlyPheGlyAspMetAsnAspSerThrLe 543  
||| : : : : :  
1246 GAC..... GACAAAGCAGCTTTTGGAAATATA..... 1272  
543 uAlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProProc 560  
||| : : : : :  
1273 .... GTTCGAGAAATTTGA..... TCCTTACCAA 1297  
560 lneuthraspleuAspPhe..... PheIleProAsnAsnLysTyr 574  
: : : : :  
1298 TCCTTCATATTTTGATTTTAAGAGATTAAATCCGTTAATAAAATCCGTT 1347  
575 GluSerLeuIleGlyThrGlnTrpHisProGluAsp..... Le 587  
: : : : :  
1348 GTAAGGCAATTTATTAACGATTTAAGGAAACACAGCTGTCCTTAT 1397  
587 uValAspIleIleArgMetGluAspLysLys..... 597  
: : : : :  
1398 AAAAGATATTCGAACACACTGAAATTAATGTAAGAAATTAACCAAGACCTA 1447  
598 ..... GluValIleProValThrHisAsnLeuThrLeuArgLysThrVal 612



```

1448 ATATCAATTTTCATACATATATGACATCATATACAGAAAATGTA 1497
      :::::::::: :::: :::::::::: ::::
613 ThrglyLeuAlaGlyAspArgThrlYsAspHehIsPheGluIleGlu 629
      :::::::::: :::: :::::::::: ::::
1498 GCATTTTCTTCT.....GAATTAACACTGATCAAGAAAT 1532
      :::::::::: :::: :::::::::: ::::
629 uLYsAsnLnLysGln.....GluLeuLeuSerGlnThrValLysThr 643
      :::::::::: :::: :::::::::: ::::
1533 AAAAGAAITTCAAATGTCACCTCAATCTATTTCTGACACTAGAAATG 1582
      :::::::::: :::: :::::::::: ::::
644 .....AspLysThrAsnLeuGluPheLysAspGly... 653
      :::::::::: :::: :::::::::: ::::
1583 CAATACATAGTCATTTGATTTAACTAAATTA.....AAGACAGTCT 1626
      :::::::::: :::: :::::::::: ::::
654 .....LysAlaThrIleAsnLeuLysHisGlyLysSerLeuThrLeuG 668
      :::::::::: :::: :::::::::: ::::
1627 ACATTTAAATTTTAAATGATCTCAACACAGAACAAATATTAACCTT 1676
      :::::::::: :::: :::::::::: ::::
668 nGlyLeuProGluGlyTyrSerTyrLeuValLysGlu.....Thra 682
      :::::::::: :::: :::::::::: ::::
1677 AGGCTTAAGTAAACCCCTTCAGTCTTAACTGAAACCAAAAGATC 1726
      :::::::::: :::: :::::::::: ::::
682 sPserGluGlyTyrLysValLysValAsnSerGlnGluValAlaAsnAla 698
      :::::::::: :::: :::::::::: ::::
1727 AAAGTCGAACCCACAAACAGATACCTTCTACTCAAAACAGAAAGCGGA 1776
      :::::::::: :::: :::::::::: ::::
699 ThrValSerLysThrGlyIleThrSerAspGluThrLeuAlaPheGluAs 715
      :::::::::: :::: :::::::::: ::::
1777 ACAAGTTCA.....ACAGATTCACAAACAAAGCTACACACTGA 1811
      :::::::::: :::: :::::::::: ::::
715 nAsnLysGluProValAlaProThr 723
      :::::::::: :::: :::::::::: ::::
1812 AAACCAAAACCGGCTGACGACAAACA 1836
      :::::::::: :::: :::::::::: ::::
seq_name: /cgn1_7/prodata/2/pna/us09_NEW_COMB.seq:us-09-430-590E-139
seq_documentation_block:
; Sequence 139, Application US/09430590E
; GENERAL INFORMATION:
; APPLICANT: Poultier, et al.
; TITLE OF INVENTION: UNUSUAL RETROTRANSPON FROM THE YEAST CANDIDA ALBICANS
; FILE REFERENCE: 674521-2001.1
; CURRENT APPLICATION NUMBER: US/09/430,590E
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,342
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 139
; LENGTH: 9850
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(9850)
; OTHER INFORMATION: 'n' can be any nucleotide 'a', 'c', 'g' or 't'
US-09-430-590E-139

alignment_scores:
  Quality: 131.50      Length: 635
  Ratio: 0.476        Gaps: 33
  Percent Similarity: 43.465      Percent Identity: 18.583

alignment_block:
US-09-494-297-2 x US-09-430-590E-139
Align seg 1/1 to: US-09-430-590E-139 from: 1 to: 9850
75 TyrGluSerTyrValArgGlyHisProTyrTyrLysGlnPheArgValAl 91
   ||| ::::::::::||| ::::::::::|||
3175 TATGGGTTTCATATCCGAAATCATACACTATATAAGAAATTT..... 3216

```

```

91 aHisAspLeuArgValAsnLeuGluGlySerArgSerTyrGlnValTyrC 108
      :::::::::: :::: :::::::::: ::::
3217 .....GATGGTCTACTCTTATGACCTTAC. 3243
      :::::::::: :::: :::::::::: ::::
108 ySPheAsnLeuLysLys.....AlaPheProLeuGlySerAsp 120
      :::::::::: :::: :::::::::: ::::
3244 ..TATGTTTATCTTAATATGATCATACCATTTTTCATTTGGAAACGAC 3291
      :::::::::: :::: :::::::::: ::::
121 SerSerValLysTyrTyrLysLysHisAspGlyIleSerThrLysPhe 137
      :::::::::: :::: :::::::::: ::::
3292 GTTTTGATA...AAATGTCATAGTCTACAGAACGCTATTTCATTAAC 3338
      :::::::::: :::: :::::::::: ::::
137 e.....GluSPyrAlaMetSerProArgLethrGly..... 148
      :::::::::: :::: :::::::::: ::::
3339 ACCATCTTCAAGAGATTAAGCTTTTCTACAGTGTATGTTGCTCTTTC 3388
      :::::::::: :::: :::::::::: ::::
149 .....AspGluLeuAsnGlnLysLeuArgAlaValMetTyr 160
      :::::::::: :::: :::::::::: ::::
3389 TCGGTACGGCTCAGATTCCTTACCTTCAGATT..TTAGTTTCCACG 3435
      :::::::::: :::: :::::::::: ::::
161 AsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluProLe 177
      :::::::::: :::: :::::::::: ::::
3436 AAAGATATCCAGTTATTAACA.....TCAACATCCGCCAAT 3476
      :::::::::: :::: :::::::::: ::::
177 uAsnAlaIleArgValIThrGlnGluAlaValIThrPyrTyrSerAspAsn 194
      :::::::::: :::: :::::::::: ::::
3477 AGCGAGATGCAAGTCAATCAATGACTATTTGGCATATATCGGGAATTA 3526
      :::::::::: :::: :::::::::: ::::
194 lAProIleSerAsnProAspGluSerPheLysArgGluSerGluSerAsn 210
      :::::::::: :::: :::::::::: ::::
3527 GCTCATTAAGCTAT..GACGATACATTC..... 3552
      :::::::::: :::: :::::::::: ::::
211 LeuValSerThrSerGlnLeuSerLeuMetArgGlnAlaLeuLysGluLe 227
      :::::::::: :::: :::::::::: ::::
3553 .....TATACCTTT 3563
      :::::::::: :::: :::::::::: ::::
227 uLLeAsProAsnLeuAlaThrLysMetProLysGlnValProAspAsp 244
      :::::::::: :::: :::::::::: ::::
3564 GAATCACCCCAATGATTCGCACAAAC..... 3588
      :::::::::: :::: :::::::::: ::::
244 heGlnLeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsnLys 260
      :::::::::: :::: :::::::::: ::::
3589 .....CAACATGATACCGTGGAGCAATATAATATC 3621
      :::::::::: :::: :::::::::: ::::
261 GlyTyrGlnAsnLeuLeuSerGlyGlyLeuValProThrLysProProThr 277
      :::::::::: :::: :::::::::: ::::
3622 GAATATGAAAC.....CGTCCAAATGT 3644
      :::::::::: :::: :::::::::: ::::
277 rProGlyAsp...ProPrometProProAsnGlnProGlnThrThrSerV 293
      :::::::::: :::: :::::::::: ::::
3645 ACCATTTGAATATCATGTGTAACCTCTCGTACAAATTCATGACGGGAA 3694
      :::::::::: :::: :::::::::: ::::
293 alLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuLeuGluGly 309
      :::::::::: :::: :::::::::: ::::
3695 TTATCGATGACACA.....GATATTAGACCTAGAGCTGATCC 3732
      :::::::::: :::: :::::::::: ::::
310 AlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgVa 326
      :::::::::: :::: :::::::::: ::::
3733 ACCTGGCAACGTATGCCGTGATGCCAACATATACATCAG...GAACAAAC 3779
      :::::::::: :::: :::::::::: ::::
326 lPheSerSerAsnAspIleGlyGluArgIleGluLeuSerAspGlyThrT 343
      :::::::::: :::: :::::::::: ::::
3780 TGTACAGACTCTGTATCAWTGAGGAGGTAGATACCATGTATCAACAGCA 3829
      :::::::::: :::: :::::::::: ::::
343 yTrThrLeu..... 345
      :::::::::: :::: :::::::::: ::::
3830 ACCAACTACCAAGATCTGGGAGGATATTAACCCGGGCAACAGGTGCC 3879
      :::::::::: :::: :::::::::: ::::
346 .....ThrgluLeuAsnSe 350
      :::::::::: :::: :::::::::: ::::
3880 ACCGATATTATTGGCAATTCAGAGATCGCGGCTTACCACTTAAACAC 3929
      :::::::::: :::: :::::::::: ::::

```



2144	ACGCGACGATCGCAAGTACATCATGACTATTGGCATATCATATCGGAGATA	2193
194	lApProIleSerAsnProAspGluSerPheIySAlrgGluSerGluSerAsn	210
2194	GCCTCAATAGACTAT...GACGCTACATATTC.....	2219
211	LeuValSerThrSerGluLeuSerLeuMetLarGlnAlaLeuIleuSgIle	227
2220	.....TTATCACCTTT	2230
227	uIleAspProAsnLeuAlaThrIySmetProIySgIlnAlProAspAsp	244
2231	GAATCACCCAAATGATTCGCACAAC.....	2255
244	heGluLeuSerIlePheGluSerGluAspIySgIyAspLySAlrgAsnIyS	260
2256	.....CAACATGATAGACCTGGAGACAATATATAAATGTC	2288
261	GlyTyrgInAsnLeuLeuSerGlyGlyLeuAlProThrIySProProth	277
2289	GAATATGAAAC.....CGTCCAAATGT	2311
277	rProGlyAsp...ProPrometProProAsnGlnProGlnThrThrSer	293
2312	ACCATTTGAATATCATGCTGAACCTCTCGTACAAATTCATCGACGGGA	2361
293	alleuIleArgIySAlrgAlaIleGlyAspTySAlrgLeuGluGly	309
2362	TTATCGATCGACCA.....GATATTAGACCTAAGCTGATCCC	2399
310	AlAThrIleuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaIrgVa	326
2400	ACCTGGCAACGATCCCGATGCCAATCCAACTACATCAG...GAACACAAC	2446
326	lPheSerSerAsnAspIleGlyGluAlrgIleGluLeuSerAspGlyThr	343
2447	TGTACGAGCTCTGTATCATGGGAGTTTGATACCATGATCAACAACGAC	2496
343	yrThrLeu.....	345
2497	ACCAATACACGATCTGGGAGGTAATACCCGGGCAACAGGTGCG	2546
346	.....ThrGluLeuAsnSe	350
2547	ACCGATATTATTGGCAATTTGAGATGCGGGCTTACACCTCTAAACAC	2598
350	rProAlaGlyTySAlrgIleAlaGluProIleThrPheIySAlrgGluAlag	367
2597	TCCGATCGATCTAGGATGACCCGAT.....GAACAG	2628
367	IySAlrgValyrThrIleIleAspGlySgIleGluAsnProAsnIyS	383
2629	ACGATATT...AGTATGACATCAAGAAATTCATGATTCGCCAAATTC	2675
384	GluIleIleValGluProTySAlrgValGluAlaIrgAsnAspPheGlu....	398
2676	GAGATGATCATATCCCATCTTACCCACA...AATGAATTGGAAACATCA	2722
399	.....GluPheSerValLeuThrThrGln.....	406
2723	AATCATATCACTTCAGGCGAGCATCTGTTATTCGAAACGAATATGAG	2772
407	.....AsnTyAlaIySPhetyTyAlaIySAsn	416
2773	CAGATACCAATTTGAAAACAAATGAAATGATTTATTCAAATTCAAAAAT	2822
417	LySAsnGlySerSerGlnValValTyrcySpheAsnAlaAspLeuLySe	433
2823	...GATGCTATTATCATTCACACACAACAATTCACCTGAAAAATTTG...TC	2866
433	rProProAspSerGluAspGlyGlyIySThrMetThrProAspPheThr	450
2867	ACATGAAAAATGGAAGAAAT.....TCATCAACATGATGAGGAACAT	2907

```

450 hrtaigluValLysTyrThrHisIleLacIlyrtaGAspleuPheIySTyr 466
2908 TGAAGACAAAAAACAACAGCGATTG.....GAATAT 2939
467 ThrValLysProArgAsp.....ThraspPr 475
2940 AATATTTCACCAACGATGAGTGTGATTAATATGACGTCAGAACGAGA 2989
475 oAspThrPheLeuLysHisIleLysValIleGluLysGlyTyrArgc 492
2990 TGACACACAAAGTCCCATGTTAAG..... 3014
492 LuLysGluIleAlaIleGluLysSerGlyLeuThrGluThrGlnLeuArg 508
3015 .....GAACCATCATATTAT.....GAACTCAAGAGTACA 3044
509 AlaAlaThrGlnLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAs 525
3045 AATGGAACAAACATGCGA..... 3062
525 pLysAspLysLeuLysAspTyrHisGlyPheGlyAspMetAsnAspSerT 542
3062 ..... 3062
542 hrLeuAlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnPro 558
3063 .....CGAATTGAAATGGCGCATTAATAGAAACCTTAAGTGTGATGGAAG 3107
559 ProGlnLeuThrAspLeuAspPhePheIleProAsnAsnAsn..... 572
3108 AATACACACAGCATTAATACGTATGTCACCTACGATTAATATATAAAAAAT 3157
573 ...LysTyrGln.SerLeuIleGlyThrGlnTrpHisProGluAspLeu 587
3158 TCAAAAGTACCAAAACAGTAAATATACGAGATCCGGAACCCAGAA..... 3201
588 ValAspIleIleArgMetGluAspLysLysGluValIleProValThrHI 604
3202 .....ACGAAATATAAAAACACACACA 3221
604 sAsn.....LeuThrLeuArgLysThrValThrG 614
3222 TTCATGTAAACCACTTAGAATTACTTGACATCAACAAATGTTTCAGA 3271
614 IlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGluLeuLys 630
3272 AGATCCTC.....AAGTTGAAGATATTGATGTCACAACTCCAAAAAAG 3315
631 AsnAsnLysGlnGluLeuLeuSer 638
3316 ACAAAATCGTTATCACCTGATTTCA 3339

seq_name: /cogn_7/ptocdata/2/pna/US09_NEW_COMB.seq:us-09-335-032-122144
seq_documentation block:
: Sequence 12214, Application us/09335032
: GENERAL INFORMATION:
: APPLICANT: Veicleulescu, Victor
: APPLICANT: Vogelstein, Bert
: APPLICANT: Kinzler, Kenneth
: TITLE OF INVENTION: Characterization of the yeast
: FILE REFERENCE: 01107/78572
: CURRENT APPLICATION NUMBER: US/09/335,032
: CURRENT FILING DATE: 1999-06-16
: PRIOR APPLICATION NUMBER: US 60/305,917
: PRIOR FILING DATE: 1997-01-23
: PRIOR APPLICATION NUMBER: US 09/012,031
: PRIOR FILING DATE: 1998-01-22
: NUMBER OF SEQ. ID NOS: 12219
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 12214

```

; LENGTH: 666448  
 ; TYPE: DNA  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-09-335-032-12214

## alignment\_scores:

Quality: 130.50 Length: 720  
 Ratio: 0.384 Gaps: 32  
 Percent Similarity: 47.222 Percent Identity: 18.333

## alignment\_block:

US-09-494-297-2 x US-09-335-032-12214 ..

Align seg 1/1 to: US-09-335-032-12214 from: 1 to: 666448

```

52 G1yLeuValG1uSerSerThrProAsnAla1leAsnProAspSerSerSe 68
    ||| :|||:||||| :|||:|||||:|||||:|||||:|||||:|||||:
296795 GGTTCATTAAGGCGCTCAATGTTCTATATGTCAGCGGAGGACACTAC 296844
    68 r.....G1uTyraGtTPTyG1yTyG1yG1uSerTyra1a 80
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
296845 TCAGCCCTTACATTAAGATTTGAAGATATGGGTATATCCACATCTAG 296894
    80 rgsly.....
296895 GTGGTGCAGAAATTACATGATATGGCAGTTGATAGAGAAATAGTTT 296944
    82 .....HisProTyTyTyLysGlnPheArgValAlaHisAspleuar 95
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
296945 TTGGAAACACACCCAGCCATAGCA.....ACTGATGAAATGCA 296982
    95 gValAsnLeuG1uG1ySerArgSerTyG1nValTyCysPheAsnLeu 112
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
296983 CGCTAATCCCAAGCCTTACCAAAATCAACCAAGCAGAGAGAGCA 297032
    112 yslYsAlaPheProLeuG1ySerAspSerSerValLysLysTPTyLys 128
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297033 AGTTAATTTAAGGCCAATTCGAGCAAGTAT..... 297067
    129 LysHisAspLysLysSerThrLysPheG1uAspTyra1aMetSerProAr 145
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297068 .....AACATAGATCATCATGATCAACGATATGATTTCCCTACTTC 297108
    145 g1leThrG1yAspG1uLeuAsnG1nLysLeuArgAlaValMetTyraSng 162
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297109 TATTAACTAGACAGGAATTCAGAGAAATTTAT..... 297139
    162 LysHisProG1nAsnAlaAsnG1yLleMetG1uG1yLeuG1uProLeuAsn 178
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297140 .....GCAGACTCGTTATTGGACATTTGCAAAACCATTAAT 297175
    179 AlalLeuArgValThrG1nG1uAlaValTPTyTySerAspAsnAlaPr 195
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297176 GAC...GCTGTACAAAACAA...TTGCGTGGCTATGTGAACAAATTTACC 297219
    195 o1leSerAsnProAspG1uSerPheLysArgG1uSerLysSerAsnLeuV 212
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297220 T.....GACATTAATGGCG 297233
    212 al.....SerThrSerG1nLeuSerLeuMetArgG1nAlaLeu 224
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297234 TCATTTTGGCGGAGGCTTCCCGAATTCACATTCGACAGATCAATTA 297283
    225 LysG1nLeu1leAspProAsnLeuAlaThLysMetProLysG1nValPr 241
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297284 ATCAACTCTATCCGAAGAA.....AAAGTCTTGAGAAATGTCAA 297324
    241 oAspAspPheG1nLeuSerLlePheG1uSerG1uAspLysG1yAspLysT 258
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297325 TCGTGAT...GAATCAGCTGTGAATGGTGTGTATGAGAGGAGTCAAGT 297371
    258 yraSnlYsG1yTyG1nAsnLeuSerG1yG1uLeuValProThLys 274
  
```

```

297372 TATCTAATCGTTAAG.....ACCAAG 297394
    |||:|||||
275 ProProThrProG1yAspProPromePro.....Pr 285
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297395 CCGTTAAATGTTGTGACCGCTCGTAAATACCTTAATTCATAAATTATC 297444
    285 oAsnG1nProG1nThrSerValLeu1leArgLysTyra1a1leG1yA 302
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297445 AAACGAATCGAACTGATATGTTTCACGCGGAGAGCTTATCCAA 297494
    302 sPTySerLysLeuLeuG1uG1yAlaThrLeuG1nLeuThrG1yAspAsn 318
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297495 ACAAAACATCTATTTCACAACACGACT.....GATTCG 297529
    319 Val...AsnSerPheG1nAlaArgValPheSerSerAsnAsp1leG1y 334
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297530 ATTCCTAATTAATTTTACCAATGATTAATTTGAGAAATGTAATTTGCA 297579
    334 uArg1leG1uLeuSerAspG1yThrTyThrLeuThrG1uLeuAsnSerP 351
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297580 AACTATACAGTTAATTCAGAGACTATAAAGATTCATATTCCTGTGATA 297629
    351 ro1aG1yTySer1leAlaG1uPro1leThrPheLysValG1uAlaG1y 367
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297630 AGTGTCTGAGAGTTCGCTATACATTAATTCGCTTCGACTGCAGAT 297679
    368 LysValTyThr1leAsp..... 374
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297680 AGATTAATCTCTATCAAGAGTTTACATTCGATTCGACAGCAAAATGA 297729
    375 .....G1yLysG1n1leG1uAsnProAsnLys..... 383
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297730 CATAGTAATCTCAGCAATTAAGAAACAAAGCAGCGCTTGGCTTTA 297779
    384 .....G1u1leValG1uProTySerValG1uAlaTyraS 395
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297780 CTTCGAGATGTTGATGATCAAAAGCCTTCTCCTCAGAACCTGCGCT 297829
    396 AspPheG1uLeuPheSerValLeuThrThrG1nAsnTyra1aLysPheTy 412
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297830 TTGCATAGCATATCAAGTGTCTCGATTAACAGATTAAGAAAGATTTCA 297879
    412 rTyra1aLysAsnLysAsnG1ySerSerG1nValTyCysPheAsnA 429
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297880 ATTCAGAAAGAAATTTAAACCTTCTGAAAGTAACCTGTATGATGCTAGAA 297929
    429 laAspLeuLysSerProProAspSerG1uAspG1yG1yThrMetThr 445
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297930 ACCTGCTAATGATGATGAATGATGAATGCAAAATGACCAAAATCCAGTA 297979
    446 ProAspPheThrThrG1uValLysTyThrHis1leAlaArgAs 462
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
297980 GAAGAGTTATCG.....GAGATGCTTAAGTAAATTTGGATTGGCTCGA 298023
    462 PleuPheLysTyThrValLysProArAsPThrAspProAspThrPheL 479
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
298024 AGATGCATCCTT.....GATCGACCTCGAGAGATATAG 298058
    479 euLysHis1le.....LysLysVal1leG1u..... 487
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
298059 TTAGCAGAAATTAAGAAATTTGAATATTAATAAAGAAATAGAACTTAC 298108
    488 .....LysG1yTyra 491
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
298109 ATGATTTCTCAAGAAACCTTGAACCTCAACATTTTAAAGAAATGCT 298158
    491 gG1uLysG1y.....G1nAla1leG1uTySerG1yLeuThrg 504
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
298159 TGAAGAAAGCCATTAAGTTACTTCAGGCTATAGAAACCCATTAACATCCG 298208
    504 LuThrG1nLeuArgAlaAlaThrG1nLeuAla1leTyTyPheThrAsp 520
    ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
  
```

```
298209 TTGAAGATTTTGGATGTCATTTGAAACCGAGTTGGCGATACCATAGAT 298258
521 SerialAgLUleuAspLysAspLysLeuLysAspLysGlyPheGlyAs 537
298259 AATGTTAGAGAGAAATTTAAAAAGATTAAAG..... 298288
537 pmetasnspSerThrLeuAlaValAlaLysIleLeuValGluTyrAlaG 554
298289 .....CACACAGCGTATGTCGTCCAGAGCGCTTATCTACATGGCAGG 298328
554 In.....AspSerAsnProGlnIleuThrAspLeuAspPhePhele 568
298329 AACCTTAACCTCTTTTAAAAATTCATTAGAGAAATAGAGAGTTCTGTG 298378
569 ProAsnAsnLysTyrGlnSerLeuIleGlyThrGlnTTPHSPProG1 585
298379 GCAAAAAAC.....CTATTGGCGAAGACCTTCGT...GA 298410
585 uAspLeuValAspLeu.....IleArgMetG 594
298411 ACATTTATTTGAAATCAATTAATTCATTTGATATGTCGACGAAACTAG 298460
594 LuAspLysLysGluValIleProValThrHisAsnLeuThrLeuArgLys 610
298461 AGGAAAAACCTGCGTTATATAA..... 298483
611 ThrValThrGlyLeuAlaGlyAspArgThr.....LysAs 622
298484 .....ACGGGTGATGAAAGTCGCTTAATGAATGAACAA 298518
622 pPheHisPhe...GluIleGluLeuLysAsnAsnLysGlnIleuLeuAs 638
298519 GTTACATTTAAGAACTTCCTCCCTACAAAGAGAGAGGAAAACTGTA 298568
638 erGlnThrVal.....LysThrAspLys 645
298569 AAGAAACCTTGACAGAGAAAAAGCAGAAACAACATGAAGAAACAAATCG 298618
646 ThrAsnLeuGluPheLysAspGlyLysAlaThrIle...AsnLeuLysH1 661
298619 ACAGTAATCAACCTGCGCTGACGATMAACTACTATTGTCAATGACAAAGAC 298668
661 sGlyGluSer 664
298669 CACCGAGTCG 298678

seq_name: /cgnl_7/prodata/2/pna/us08_NEW_COMB.seq:us-08-956-171C-268

seq_documentation_block:
; Sequence 268, Application US/08956171C
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; G11 H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956, 171C
```

```

; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoover, Kenley K.
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 610-5790
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 268:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4702 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 268:
US-08-956-171C-268

alignment_scores:
Quality: 129.50 Length: 722
Ratio: 0.363 Gaps: 31
Percent Similarity: 49.446 Percent Identity: 18.837

alignment_block:
US-09-494-297-2 x US-08-956-171C-268/rev ..
Align seg 1/1 to reverse of: US-08-956-171C-268 from: 1 to: 4702

14 AsnThrGlnArgValLeuSerLysAsnSerLysArgPheThrValThrLe 30
2170 AACACAGAAAAAGTTGTAAGTAATGTACTTAACGACCTTACAAAAACTT 2121
30 uValGlyValPheLeuMetIlePheAlaLeuValThrSerMetValGlyA 47
2120 AGTCGAAATTTATTTAAACA.....ACGATCATGTGTAATTTACATTAATG 2077
47 lAlaYstHValPheGlyLeu.....ValGluSer 56
2076 CACAAAAAATGTTGGCGCTATTATGACGCGTGAACATGCTGTGATAGT 2027
57 SerThrProAsn...AlaIleAsnProAspSerSerSerGluTyrArgTr 72
2026 AATTCCTCGAATTTACTTATTAATCA..... 2000
72 pTyrGlyTyrGluSerTyrValArgGlyHisProTyrTyrLysGlnPhea 89
1999 .....ATTACGACTTCCCGGAATTAATTATACAGATA 1969
89 rGValAlaHisAspLeuArgValAsnLeuGluLysSerArgSerTyrGln 105
1968 CGCTTGTAATTCATTTCTGCACAAACAAACATTACAAAATGCTTCCAA 1919
106 ValTyrCysPheAsnLeuLysAla.....PheProLeuG1 118
1918 ACATTCATAATATCTTTATTGAGTCGGAATTCAGATACGTTCAAGGTAA 1869
118 ySerAspSerSerValLysLysTyrPtyrLysHisAspGlyLe... 133
1868 CACAGATATATAGTTGTCGACTTATATGAAACAAATCATATATTGTG 1819
133 ..... 133
1818 ACGAGCAACAATACAGCGATGATAAATGTTACAGATTAATTAATCGCAA 1769
134 .....SerThrLysPheGluAspTyrAlaMetSerProArgIleThrG1 148
1768 AAGATACGCTGGAACCTGATACATATATCAATGATGATTAATAACAGATGA 1719
```

```

148 yaspgluleuasnlglnlysluvalglalvalmettyrasnlglyhispro 165
1718 cagcccaattgatacaca...tcagatgacagatgacagtaagaag 1672
165 lnsnlaasnlglyilemetgluglyleugluProleuasnalailearg 181
1671 aatataaacaactgttaagaaactgtatgaatgaagaaacatt 1622
182 valthrglnlualavaltrpytyrtyrseraspasnalaProlesers 198
1621 caatcacaagatcaccattt.....tcaaaagtrtgattgaaga 1581
198 nproasgluserphelysargluserguserasnleuvalserthrs 215
1580 nrttccttaagcaattaacagaaatcactgcagatgacgttgcataaaca 1531
215 ergluserSerleuMetarglnalaleuylsglnleuileasproasn 231
1530 aagacttacaaagatgcgctaaatagcattaaa...atgaacaatgcctcaa 1484
232 leuvalathrllysmetprolysglnvalProaspaspheglnleuserl 248
1483 ttccgctgaatacttagaagaacacttcattgatgat.....at 1446
248 epheglsusergluaplysglyasplys.....tyrasnlysglytyrg 263
1445 tgcacaaagaaacctgatacagatacacaatttatctatnaacatgctaac 1396
263 lnsnleuuserSerlglyleuvalProthrlys.....Pro 275
1395 aagactttatagctgacgcttaaatgacagatgaactaataaatacga 1346
276 prothrProglaspprometProproasnlgProglinthrse 292
1345 gcaattgctcaaaagacaaacgttataaaatgaatataatttgaaaaa 1296
292 valleuilearglystyrAlaleglyaspyrtyrserlysluenglug 309
1295 accgttaagcaaacacatttaatttaacagattgacatnaacacagttcgcc 1246
309 lvalathrlleuThrglyAspasnValasnserPheglnalilearg 325
1245 aagacacacaaagttgatattgaatgatgctcaaaagcacaacgtacgaa 1196
326 valPheSerSerasnspIlelglyluarglleuIuseraspglyth 342
1195 aagatttaaaagtatgatatttaacaa..... 1169
342 rtyrThrlleuThrgluuasnserProalaglytyrserIlealaglup 359
1168 .....ttaaactgttcacacagatcct..... 1148
359 roilerhrPheIysValgluAla.....glyIysValtyr 370
1147 .....catttttaattttgaagcgacacatttaaaatttaattgtaaaataat 1103
371 ThrIlelleasp...glyIysglnIlegluasnProasnlysgluileva 386
1102 gacatttaagagatcaaaagtttcaactgatacatttaacaaagaa..... 1058
386 lgluprotyrserValglualatyrasnaspheglnlupheserVal 403
1057 .....tataaagttgacatcatt.....ggcgttg 1033
403 eutThrThrglnasnTrAlalysPheTyrtYrAlalysAsnlySAsngly 419
1032 ctaaatgtgaaaaagatgctgagaaagatttcttaaaagatnaaaacatg 983
420 serSerGlnIvalValtyrcysPheasnalaSpleuylsserProproas 436
982 catttcaaatgttatttt...ggacaaagcaaatgctcaagatgaaccnaa 936

```

```

436 psergluaspglyglylysthrmetthrProaspPheThrThrglyglu 453
935 T.....GATAAGAAAGCAACGAGTGTGTGATGTAACATGAATCAT 893
453 allystyrThrhIsIlealagIyArGaspLeuPheIysTyrtYrValys 469
892 .....AACCTGATGGTGC...TTATCGAAAGATGCAATTAAAC 857
470 ProargaspThrAspProaspThrPheleuIysHisIleIysValIl 486
856 caccatvtgagtcattatctagcttgcattgacattatnaaaatgacac 807
486 eglu...lysGlyTyrtArgGluIys..... 493
806 agatcacaagacgacgaagatnaaacattgcacaaacgtttaattg 757
494 .....glylnAlaIlegluTyrseryleuThrglnThrgln 506
756 atatgatgcttgacaaagtatc.....aatgcacatgaa 722
507 leuargAlaIathrglnleuAlaIleTyrtYrPheThrAspserAlagl 523
721 agttcacaagacgataaagtagctgctgattacatcaaatgattcattgca 672
523 uluasplysAsplysleu..... 529
671 agaaactcagacaaactgatgatgacatttttaatacaaaaagaata 622
530 ....lysaspTyrhIsglyPheGlyAspMetasnspserThrleuAla 544
621 caacaaataatmaagaaatatttccaaagctgattgattcagttagaanaac 572
545 valAlaIysIleleuvalgluTyrtAlaglnaspserasnProProgluile 561
571 gtttaaaagacttttgctgaagaccacagaacaaataatgatgaag 522
561 uthraspLeuaspPheIleProasnasn..... 572
521 caaaaatgatgaatttaataatgcattgcttcaaaatttagatnaaaattta 472
573 .....lystYrGlnserleuIlegluThrglnThrhIsPro 584
471 gtaaatatttgaaaagacgacgaattgctatcagatnaacaaagatca 422
585 gluaspLeuValaspIleIleargmetgluasplyslysgluvalIlepr 601
421 aaatcaattgcacatctggt.....AGTGGCAATTAATCA 384
601 ovalThrhIsasnleuThrlleuarglysthrValThrglyleuAla...g 617
383 agtcgacaaatnaatgtgat...aagctacatgcgacagctgcacattag 337
617 lysaspargThrlYsaspPheHisPhegluIlegluleuIysAsnAsnlyS 633
336 gcgttaagccttaacgatttgaatcgtcaaatggcttaaaacgataaagat 287
634 glnIgluleuIuserGlnThrValys..... 642
286 aatgagttgttcgctaaagaatttaaaagatttaacaaattctaaagat 237
643 ....ThaspIysThr 646
236 gccgcacagacaaamaac 221

```

seq\_name: /cgcn1\_7/ptodatta/2/pna/US08\_NEW\_COMB.seq:US-08-956-171C-264

```

seq_documentation_block:
; Sequence 264, Application US/08956171C
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen

```



```

2074 AATGTTATGTTGGTGGACAAAGTATATCATCAAGTCAACAGCTGATTAAT 2025
381 ProAsnLysGluIleValGluProTyrSerValGluAlaTyrAsnAspH 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2024 GCTGGCAAGTTAAATTTTCAACCTCAAAAT.....AACAGCTT 1987
397 eglGluPheSerVal.....LeuThrThrGlnA 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1986 CCAAGATTTTGAATATTAAGTATATGATGATCAAGATGATGATGATGATCA 1937
407 sntYrAlaLysPheTyrTyrAlaLysAsn.....Lys 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1936 AATATGCAAGTCAAAACATGACACGTAATATTTTCAAGATGATGATGCGAAA 1887
418 AsnGlySerSerGlnValValTyrGlyPheAsnAlaAspLeuLysSerPr 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1886 AGGGTACGACCAACTTTTCATATATCAATGACAGCC..... 1851
434 oProAspSerGluAspGlyGlyLysThrMetThrProAspPheThrThrg 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1850 .....TCACACAGTGGCGCGCAAAATTTTACAACAAGTCAATTTTG 1811
451 LysGluValLysTyrThrHisIleAla..... 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1810 GAACATTCGAATATACAGAGTCTGCTGNTACACAAGTGAAGATGCTTGAT 1761
460 .....GlyArgAspLeuPheLysTyrThrValLysProArgAspTh 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1760 GTAACAACAGGTAAAGATATATTT..... 1737
473 rAspProAspThrPheLeuLysHisIleLysLysValIle..... 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1736 .CCACCACAAACATATTCAGCAAAATGTTGATCAAGTGTACATGATGATA 1688
487 .....GluLysGlyTyrArg.....Glu 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1687 ATCAGCAATCTGCATTGACTGCTAAAGATATTAACATCACTACGCTCGTGAT 1638
493 LysGlyGlnAlaIleGluTyrSerGlyLeuThrGluThr...GlnLeuAr 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1637 AGTTTCATTCGCTCAACTTATTAATGATACAAATRAAACTGTAATAATGAC 1588
508 gAlaAlaThrGlnLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuA 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1587 GAATGCTGGACAAATCAAGTACATATATTTTAC..... 1554
525 sPlsAspLysLeuLysAspTyrHisGlyPheGlyAspMetAsnAspSer 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1553 .....GATGTAAAGACACCA 1539
542 ThrLeuAlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnPr 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1538 ACTGTAACTGTAGCAATCAACCATAGAAAGTGGTAAACATGAATGCC 1489
558 oProGlnLeuThrAspLeuAsp..... 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1488 TATTTGATTTGACTACAAAGGATATGTTAGTGGAGTGTGACAAATACAG 1439
566 .PhePheIleProAsnAsnAsnLysTyr.....GlnSerLeu 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1438 TTACAGGATTTACCAAGCGATTAAGTTAGTACATAGTCAACGAATTCATC 1389
578 IlleGlyThrGlnThrPheHis...ProGluAspLeuValAspIleIleArgMe 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1388 ATTGGGACACCAACAAAATTTGGTCAATCAACAGTGCACGTGTGTCTAC 1339
593 tGlu.....AspLysLysGluValIleProValThrHisAsnLeu... 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1338 TGACCAAGCAAAATACAAATGACAGACAACTTTTACATAATATGTGTGG 1289
607 .ThrLeuAlaGlyLysThrValThrGlyLeuAlaGlyAspArgThrLysAsp 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1288 ATACGACAGCACCAACAGTGCACCAATA...GGAGATCAATCATCAGAA 1242

```

```

623 PheHisPheGluIleGluLeuLysAsnAsnLysGlnGluLeuLeuSerL 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1241 GGTATTCACCAATATATCCCC..... 1222
639 ntThrValLysThrAspLysThrAsnLeuGluPheLysAspGlyLysAla 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1221 .....GATTAAANATTTTACGCAAGATTAACAGTGAAGAAATGCCG 1185
656 hrIleAsnLeuLysHisGlyGluSerLeuThrLeuGlnGlyLeuProGlu 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1184 TGACGAAT.....ACAGTGAAGTGAATGCCATGCCATCC 1156
673 GlyTyrSerTyrLeuValLysGluThrAspSerGluGlyTyrLysValL 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1155 GCACATAACATTTGATAGTACAAATATATCTATATGATGTTGTTACACCA 1106
689 sValAsnSerGlnGluValAlaAsnAlaThrValSerLysThrGlyIle 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1105 CATTGTACAAAGTACTATATCAATTCGTTCTTACAGATGCCAGCGGTAA 1056
706 hrSerAspGluThrLeuAlaPheGlu 714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1055 AAACGACGACACTTTTAAATATGAA 1030

```

```

seq_name: /cgn1_7/ptodata/2/pna/US08_NEW_COMB.seq: US-08-956-171C-168

```

```

seq_documentation_block:

```

```

; Sequence 168, Application US/08956171C

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Charles Kunsch

```

```

; GILL H. Choi

```

```

; Patrick S. Dillon

```

```

; Craig A. Rosen

```

```

; Steven C. Barash

```

```

; Michael R. Rannon

```

```

; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

```

```

; NUMBER OF SEQUENCES: 5235

```

```

; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Human Genome Sciences, Inc.

```

```

; STREET: 9410 Key West Avenue

```

```

; CITY: Rockville

```

```

; STATE: Maryland

```

```

; COUNTRY: USA

```

```

; ZIP: 20850

```

```

; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

```

```

; COMPUTER: HP Vectra 486/33

```

```

; OPERATING SYSTEM: MSDOS version 6.2

```

```

; SOFTWARE: ASCII Text

```

```

; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/956,171C

```

```

; FILING DATE: 20-Oct-1997

```

```

; CLASSIFICATION: <Unknown>

```

```

; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 60/009,861

```

```

; FILING DATE: January 5, 1996

```

```

; APPLICATION NUMBER: 08/781,986

```

```

; FILING DATE: January 3, 1997

```

```

; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Hoover, Kenley K.

```

```

; REGISTRATION NUMBER: 40,302

```

```

; REFERENCE/DOCKET NUMBER: PB248P1

```

```

; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (301) 610-5790

```

```

; TELEFAX: (301) 309-8439

```

```

; INFORMATION FOR SEQ ID NO: 168:

```

```

; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 7963 base pairs

```

```

; TYPE: nucleic acid

```

```

; STRANDEDNESS: double

```

```

; TOPOLOGY: linear

```

```

; SEQUENCE DESCRIPTION: SEQ ID NO: 168:

```



US-08-956-171C-168

alignment\_scores:

Quality: 128.50 Length: 734  
 Ratio: 0.377 Gaps: 37  
 Percent Similarity: 46.458 Percent Identity: 18.801

alignment\_block:

US-09-494-297-2 x US-08-956-171C-168/rev ..

Align seg 1/1 to reverse of: US-08-956-171C-168 from: 1 to: 7963

```

101 SerArgSerTyrGlnValTyrCysPheAsnLeuLysLysAlaPheProLe 117
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7385 ACAAAATCCAAATTTGTTAT...TTCAACAATAATTTCTTATTC.... 7344
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 uGlySerAspSerSerValLysLysTyrPtyrLysLysHisAspGlyLys 134
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7343 .....CACGATGCAACAATGCTTGAAMTCACCTGACAGACAGCA 7304
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 eRhrLysPheGlnAspTyrAlaMetSerProArgLLeRhrGlyAspGlu 150
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7303 ATGAAAAATGCGTATTTATGCTGAAAAATACCAAACTGTTACAAATGAA 7254
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 LeuAsnGlnLysLeuArgAlaValMetLysnGlnHisProGlnAsnAl 167
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7253 TTTAAATCAT...TTAAATGCTCAATTA.....GAAGAAAAATTA 7219
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 aAsnGlyLeuMetGlnGluLeuGluProLeuAsnAlaIleArgValTArg 184
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7218 TAAAGAATTTACAGACATTCGATTTTATCAATGTT..... 7182
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 InGlnAlaValTPTyTyrLysSerAspAsnAlaProIleSerAsnProsp 200
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7181 .....GGTACTGAA 7173
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 GLUserPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLe 217
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7172 GAAGATTTTTCACCAATCATCAGAAAGATTACCAAACTTATCTAATAATT 7123
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 uSerLeuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeuAlaT 234
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7122 GAGTCGTTTATGATTTAACTAAATATCTTGAGACCAAAACATATCTT 7073
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 hRlysMetProLysGlnValProAspAspPheGlnLeuSerIlePheGlu 250
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7072 ATGATTTAAGTTCTAGTTTAACTGAAAAAACAACTGCACAACTT...GAA 7026
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 SerGluAsp.....LysGlyAspLysTyrAsnLysGlyTyr 262
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7025 GAAGAAATCATTTATTTGCTACTCAAGTTGACGAATATATGAGCAATA 6976
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
262 rGlnAsnLeuLeuSerGlyGlyLeuValProThrLysProThrProG 279
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6975 TCTTGAATGCAACGA.....CAAGTCA 6953
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
279 LysAspProPheMetProAsnGlnProGlnThrSerValLeuIle 295
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6952 GTGATTTAAGTCACAAATCAATCAATGCAAACTGATACAACTTCTCT 6903
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 Arg.....LysTyrAlaIleGlyAspTyrSe 304
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6902 AATTTAAGACATGAAATATCACAGTCTTAAAAATCAACTTAATGATATGCG 6853
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 rLysLeuLeuGlnGlyAlaThrLeu.....GlnLeuThrGlyA 317
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6852 AAAA.....GATGGCGCAAGTTAAGTATTTACAAAGTTTAGTGATG 6809
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
317 sPasnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleGly 333
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6808 AACCATTAACAAATTAAGATAAACGTTTG...CCTCAAGTTATTAAT 6762

```

```

334 GluArgIleGlu.....LeuSerAspGlyThrTyrThrLeuTh 346
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6761 GAAGCGGTAGAAATTTAGACATTTAAACAGATGCGACAGATATGATAT 6712
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 rGlnLeuAsnSerProAlaGlyTyrSerIleAlaGluProIleThrPhe 363
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6711 TAACATTAATGAA.....GATTCATTAATGCGGTTA 6683
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
363 ySValGlnAlaGlyLysValTyrThrIleIleAspGlyLysGlnIleGlu 379
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6682 AACATGTTAATGGTCATTAATAT.....GATCCTGTTGAAGTAACTAG 6642
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
380 AsnProAsnLysGluIleVal..... 386
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6641 CAATCTACAAAGAAATTAATTCTTAACTAGCTTACGTATCACTTAATTA 6592
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 .....GluProTyrSerValGlnAlaTyrAsnAspPheG 398
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6591 GGTACTAGACCATATTTATCGTCCCATTAATGTTGATGATGATGATG 6542
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398 InGluPheSerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAla 414
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6541 TTCATTTTGAATAAAAAACGTACTGAAAAAATGTAATTAATTATGATCA 6492
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415 LysAsnLysAsnGlySerSerGlnValAl.....Ty 425
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6491 TTATCAGAACACTA...TCAAGTACTTACTTTCATGCTGTAAGATAATA 6443
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
425 rCysPhe..... 427
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6442 TTGTTCCATCAAAAAGAGTATTAATTAACAAATAGAGGAAGCGCGG 6393
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
428 .....AsnAlaAspLeuLysSerProProAspSerGluAsp... 439
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6392 AAACGATGAGAAATATAGAGATCTAAATCCCGGAGATTCAGTTATC 6343
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
440 .....GlyGlyLysThrMetThrProAs 447
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6342 TTTTCTTAGTCATAAAGCTACACAGGGGTGTAACAGCAAGTAAGA 6293
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
447 pPheThrThr.....GlyGluVal.....LysT 455
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6292 TTATATGACATTTACATTTGCAAGATTAAGTGGTGAATTAAGGGAAT 6243
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
455 yThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysProArg 471
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6242 TTTGGACGCTACAAAAATGATATG.....GCAACATCAAGCCTGAA 6199
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
472 AspThrAspProAspThrPheLeuLysHisIleLysLysValIleGlyL 488
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6198 GAA.....ATTGTCATGTT..... 6184
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
488 sGlyTyrArgGluLysGlyGlnAlaIleGluTyrSerGly.....L 502
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6183 .....AAAGTGACATCATTAACATATCGCGGAATAAACA 6147
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
502 eutRhrGluThrGlnLeuArgAlaAlaThr.....GlnLeuAlaIle 515
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6146 TGAAGTCACAAATTAATGACTAGCGACAACTGAAGATTAATTAACA 6097
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516 TyrTyrPheThrAspSerAlaGlnLeuAspLysAspLysLysAsp... 531
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6096 GAACAATTTGATGATGTCACCTTTATCACCGGACAGAAATACAAAGA 6047
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
532 ....TyrHisGlyPheGlyAspMetAsnAspSerThrLeu...AlaValA 546
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6046 GATTTCCATTAATTTGTCAGATATTTGAATAATGTAATTTACAGATATCA 5997
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
546 lAlysIleLeuValGluTyrAlaGlnAspSerAsnProProGlnLeuThr 562
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5996 CACGTCATTTATGTAATAAATAATATCAAGAACGA..... 5965
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
563 AspLeuAspPhePhe.....IleProAsnAsnAsnLysTyrGlnSe 576

```

```

5964 .....TTTACACATATCCAGCTGCTAGTTCATCATCATTAACCT 5924
5965 rleuilegylthrclothrphs.....ProGlu 586
5923 TCGCAGTGGCTTACGTATCATCTATTACGAGTGTACGATTGCCAATAAT 5874
586 spleuValaspleilearMetGluAspLysGluValle..... 600
5873 CAATTGTGACATTTATCCATTTGTAACAAAGTTGTATATATAGTGT 5824
601 .....ProVal1th 603
5823 ATTATTTCGATGATATTGTTAAAGTTAGAAATGAGTCTGCTGTGC 5774
603 rHAsnLeuThleuArGlyThrValThnGly.....Leu1ag 617
5773 GACGTCTATACAGTCCAGTAACCTATTATAGCACATCTCCGATTGCCA 5724
617 LysAspArGlyThrLysAspPheHisPheGluileGluLeuLysAsnLys 633
5723 GTGATGAAAGTAGTGAAGCAGCTCGTAATGAAACATTGAAGAGAGAA 5674
634 GluGluLeuLeuSerGlnThrValLysThrAspLysThrAsnLeuGluPh 650
5673 ATCATGTTGTTAGACATATGATTTATCTCATCATCATGTAAGTAGAGTA 5624
650 eLysAspGlyLysAlaThrIleAsnLeuLysHisGlyGluSerLeuThrL 667
5623 TGGTTCCTCCAAACTGCCATAC...TTAAAGAGCAGAAATTTTA.... 5561
667 euGlnGlyLeuProGluGlyLysSerTyrlleuValLysGluThrAspSer 683
5580 .....TGCATATATC..... 5572
684 GluGlyTyrlLysValLysValAsnSerGlnGluValAlaAsnAlaThrVa 700
5571 GATTAATTCATGCTAGAAATGAAATATGTTGAAAAAGCATATAAAAAAC 5522
700 lserLysThrGlyLthrSerAspGluThrLeuAlaPheGluAsnLysNL 717
5521 TGCACAG...GGTCAGTTTACAGATAAATATTTGGCTTGAAATGCTGA 5475
717 ys 717
5474 GA 5473
seq_name: /cgnl_7/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-03782A-2
seq_documentation_block:
; Sequence 2, Application PC/TUS0103782A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Ford, John E et al
; TITLE OF INVENTION: Novel Bone Marrow Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-040
; CURRENT APPLICATION NUMBER: PCT/US01/03782A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/250,583
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: Custom
; SEQ ID NO 2
; LENGTH: 16122
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-03782A-2

```

```

alignment_scores:
  Quality: 127.00      Length: 722
  Ratio: 0.372        Gaps: 36
  Percent Similarity: 47.230  Percent Identity: 20.637

alignment_block:
US-09-494-297-2 x PCT-US01-03782A-2 ..
Align seg 1/1 to: PCT-US01-03782A-2 from: 1 to: 16122

84 TyrTyrlLysGlnPheArGValAlaHisAspLeuArGValAsnLeuGlu1 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6967 TATGAGAACTAGAGGGGAGTACTCTCATGAACGCCGAGAAAGCCCTTCAGGC 7016
100 y.....SerArgSerTyrlGlnValTyrlCysPheAsnLeuLysLys1ap 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
7017 TATCTTCACAGAAATGAGAGAGGTCCAC.....AAGAGGCA. 7053
115 heProLeuGlySerAspSerSerValLysLysTrpTyrlLysHisAsp 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
7054 .....ACTCTGCTGCAGTGCAGTGCATCAACAAAGAG 7086
132 GlyLLeSerThrLysPheGluAspTyrlAlaMetSerProArGylThrG1 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
7087 GAAGTC...CTGAATCCATGATGCCATGTCATCTCCAAACAAAGACAGA 7133
148 yAsp.....GluLeuAsnGlnLysLeuArGAlaValMet 160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7134 AACAGTGAAGCCCAACAGTGAATCTCAAGGCCCTTCTGCTGAGTGTG. 7182
160 yRAsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluPro 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7183 .....GAACACAAATTCCTCAAAATTCAAAAAGTAAGAAAGCC 7221
177 LeuAsnAlaIleArGValThnGlnAlaValTrpTyrlSerAspAs 193
   ||| : : : : : : : : : : : : : : : : : : : : : : : : :
7222 CTGCTGATTAATCTGTCACATAT..... 7245
193 nAlaProIleSerAsnProAspGluSerPheLysArg..... 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7246 ...CCCACCTCACAGAGACAGAAATGTGAAGAAATTCAGAGAGAAC 7291
205 ..... 205
7292 TCAATTCGCCATGGGAAGGCCACCTGAGTTACTGTGCTCGGCACAAAG 7341
206 .....GluSerGluSerAsnLeu.....ValSerThrSe 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7342 CAGCTAGAGGAATCTGCAGATCATCTGCTTCCAGGCTGCAGAAATC 7391
215 rGlnLeu.....SerLeuMetArgGlnAlaLeu 225
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7392 CCACCTCCGCCCGGTGCTGATGAGAAAGAACATGATAGCGAGTGTGG 7441
225 yGlnLeu.....IleAspProAsnLeuAlaThrLysMetProLysGlnVal 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7442 GGGCCCTGTCTATTGACCCCAACATGTTGAATGCACAAAGCAACAGCTC 7491
7492 .....CAGTTATGCTAAAGAAATTTGAAGCAGCAGCAACGACATGA 7535
241 ProAspAspPheGlnLeuSerIlePheGluSer...GluAspLysGlyAs 256
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
7536 GCAACTGAATGAGCAGCTCAGGCATCTTAACAGGC..... 7572
256 pLysTyrlAsnLysGlyTyrlGlnAsnLeuLeuSerGlyLysValProT 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
273 hrLysProProThrProGlyAspProPrometProProAsnGlnProGln 289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7573 .....CTGAGAGATCTCTCTGTCACACAGCAAGTACAG 7608
290 ThrThrSerValleuIleArGlyLysTyrlAlaIleGlyLysPyrSerLysLe 306

```

```

7609 AAAGACTCCAGACATCAATCAG..... 7632
      |||:||||:
306 uleugluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheG 323
      |||:||||| |||:||||| |||:|||||
7633 .....AATGGTGTGAGTACT...GACAACTCAACTACCTCCGCT 7669
      |||:||||| |||:||||| |||:|||||
323 LnaIaArgValPheSerSerAsnAspIleGlyGluArgIleGluLeuSer 339
      |||:||||| |||:||||| |||:|||||
7670 CCAGCCCAAAAT.....GACCAAGCTATTGTAAAGACACCCAGTACACG 7713
      |||:||||| |||:||||| |||:|||||
340 AspGlyThrTyThrLeuThrGluLeuAsnSerProAlaGly.....Ty 354
      |||:||||| |||:||||| |||:|||||
7714 GAAGTGTCCCGAGACTTATCAGAAAGGTGAGGCGAGTGGACAAACGGCT 7763
      |||:||||| |||:||||| |||:|||||
354 rSerIleAlaGluProIleThrPheLysValGluAlaGlyLysValTyT 371
      |||:||||| |||:||||| |||:|||||
7764 GAGTGTCCAGTCACTATACAGCACCAACGAGGCTGTAAAG..... 7806
      |||:||||| |||:||||| |||:|||||
371 hrIleIleAspGlyLysGlnIleGluAsnProAsnLys..... 383
      |||:||||| |||:||||| |||:|||||
7807 .....CAGCAATTTGAGAGAGACCACTGAATTCGATCTGAC 7842
      |||:||||| |||:||||| |||:|||||
384 ...GluIleValGluProTySerValGluAlaTyAsnAspPheGluI 399
      |||:||||| |||:||||| |||:|||||
7843 TTGGAGCAGTTAGACACAGAGTTAAGAGGCTCAGACACTGTGCGATGA 7892
      |||:||||| |||:||||| |||:|||||
399 uPheSerValLeuThrThrGlnAsnTyAlaLysPheTyTyAlaLysA 416
      |||:||||| |||:||||| |||:|||||
7893 ACTCTCAGTGTCTATGTGTGAGCAGTCACTCAAGATGAACTAAGAAG 7942
      |||:||||| |||:||||| |||:|||||
416 sn.....LysAsnGlySerSer 421
      |||:||||| |||:||||| |||:|||||
7943 GTTTGGAGACAGTTGCCCTGCTTCCAAAGTTTAAAGACCTTGACGCC 7992
      |||:||||| |||:||||| |||:|||||
422 GlnValValTyTyCysPheAsnAlaAspLeuLysSerProProAspSer 438
      |||:||||| |||:||||| |||:|||||
7993 GATCGCACTTAACAGACTCCAGGCACTTGTCCAGCACCCAGCGATTCCA 8042
      |||:||||| |||:||||| |||:|||||
438 u.....AspGlyLysTyThrMetThrProAspPheThrThrGlyG 452
      |||:||||| |||:||||| |||:|||||
8043 GCAAATGTTTATGATGAGTGAAGACCTGTTGATGAATAACAAGCCGCG 8092
      |||:||||| |||:||||| |||:|||||
452 LuValLysTyTyThrHisIleAlaGlyArgAspLeuPheLysTyThrVal 468
      |||:||||| |||:||||| |||:|||||
8093 AACCAAAAACCTGCCCAATTTCTGCAAAA..... 8121
      |||:||||| |||:||||| |||:|||||
469 LysProAlaArgAspThrAspProAspThrPheLeuLysHisIleLysVal 485
      |||:||||| |||:||||| |||:|||||
8122 .....TTGGAGCGGGCTACAGTCCA 8141
      |||:||||| |||:||||| |||:|||||
485 LileGluLysGlyTyArgGluLysGlyGlnAlaIleGluTySerGlyL 502
      |||:||||| |||:||||| |||:|||||
8142 GCTACAGAGAGAATGAGAGTTTCAGAAAAGCTTAATCAACACAGTGGCT 8191
      |||:||||| |||:||||| |||:|||||
502 euThrGluThrGlnLeuArgAlaAlaThrGlnLeuAlaIle-TyTyTyPh 518
      |||:||||| |||:||||| |||:|||||
8192 CCTATGAGGTCAATGTGCTCAAGGGGGAATCTCACTTCTTCTGACT 8241
      |||:||||| |||:||||| |||:|||||
518 eThrAspSerIleGluLeuAspLysAspLysLeuLysAspTyThrHisGlyP 535
      |||:||||| |||:||||| |||:|||||
8242 CCTG.....GAGAAAGAGAAAAGAGCTTACAAA 8270
      |||:||||| |||:||||| |||:|||||
535 heGlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGlu 551
      |||:||||| |||:||||| |||:|||||
8271 CCAGTGTGTTGAGCTCAAAAAC.....ATTGGGAAGAGCTTAGTAA 8314
      |||:||||| |||:||||| |||:|||||
552 TyAlaGlnAspSerAsnProProGlnLeuThrAspLeuAspPhePheI 568
      |||:||||| |||:||||| |||:|||||
8315 AAAGTGCAGACAGCAATCCAG.....CTCAAGATTTGTAT 8351
      |||:||||| |||:||||| |||:|||||
568 eProAsnAsnAsnLysTyGlnSerLeuIleGlyThrGlnThrPheProG 585
      |||:||||| |||:||||| |||:|||||

```

```

8352 GCAGAAAGCTCAAAATAT.....CAGTGGCATGTGG 8383
      |||:||||| |||:||||| |||:|||||
585 LuAspLeuValAspIleIle.....ArgMetGluAspLys 596
      |||:||||| |||:||||| |||:|||||
8384 AAGACCTTGTGTCATGATGATGAAAGATTGTAAAGCTTAAGATGTCTGACTGG 8433
      |||:||||| |||:||||| |||:|||||
597 LysGluValIle...ProValThrHisAsnLeuThrLeuArgLysTyIva 612
      |||:||||| |||:||||| |||:|||||
8434 CGAGTCACTGTGATCCAGTCCAGTCACTAGATCCAGTCTCTTAAGATCAA 8483
      |||:||||| |||:||||| |||:|||||
612 LThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGlu 629
      |||:||||| |||:||||| |||:|||||
8484 GCGTATGCTG.....AATGAGTGAAGAACGGCGCTCCCTGCTGGAAA 8527
      |||:||||| |||:||||| |||:|||||
629 euLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysTyAspLys 645
      |||:||||| |||:||||| |||:|||||
8528 TATGAAATAGTGTCTGACATCTGATCAATTCCTCAAGAACGATGAG 8577
      |||:||||| |||:||||| |||:|||||
646 ThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsn..... 658
      |||:||||| |||:||||| |||:|||||
8578 GATGGAATC.....CGGATGAGAGAGGCTGGGATCAACAGAACATGGA 8621
      |||:||||| |||:||||| |||:|||||
659 .....LeuLysHisGly.....GluSerLeuT 666
      |||:||||| |||:||||| |||:|||||
8622 TGCCTTACAGAAAGCTGCAGGCCCAAAACAGGCTCACTCGAATAATGA 8671
      |||:||||| |||:||||| |||:|||||
666 hr.....LeuGlnGlyLeuProGluGlyTyTySerTyThrValLysGlu 680
      |||:||||| |||:||||| |||:|||||
8672 CTCAGAGCTCTCAGGAGAGTCCAGAAAGCTTAAATATTAAGAAAGAG 8721
      |||:||||| |||:||||| |||:|||||
681 ThrAspSerGluGlyTyTyLysValLysVal.....AsnSerG 693
      |||:||||| |||:||||| |||:|||||
8722 GTTGAAGAGGCCAAACCAACTTGAGATCTTGAATGCTGTGGTTCCA 8771
      |||:||||| |||:||||| |||:|||||
693 nGluValAlaAsnAlaThrValSerLysTyThrGlyIleThrSerAspGlu 710
      |||:||||| |||:||||| |||:|||||
8772 AGCCTGTAACAAAGAACTGGAAGAACTTAAGACTCAACAGGAAGTGC 8821
      |||:||||| |||:||||| |||:|||||
710 hrLeuAlaPheGlu 714
      |||:||||| |||:||||| |||:|||||
8822 TGCAGGCCCTGAGAG 8835
      |||:||||| |||:||||| |||:|||||

```

seq\_name: /cgnl\_7/prodate/2/pna/PCT\_NEW\_COMB.seq:PCT-US01-03782A-190

seq\_documentation\_block:

```

; Sequence 190, Application PC/TUS0103782A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Bone Marrow Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-040
; CURRENT APPLICATION NUMBER: PCT/US01/03782A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/250,583
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: Custom
; SEQ ID NO 190
; LENGTH: 16341
; TYPE: DNA
; ORGANISM: Homo sapiens
; PCT-US01-03782A-190

```

alignment\_scores:

```

Quality: 127.00      Length: 722
Ratio: 0.372         Gaps: 36

```

Percent Similarity: 47.230 Percent Identity: 20.637

Alignment block:  
US-09-494-297-2 x PCT-US01-03782A-190 ..

Align seg 1/1 to: PCT-US01-03782A-190 from: 1 to: 16341

```

      84 TyrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGluG1 100
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      7186 TATGAGAAACTAGGGGAGTACTCTCATGACGCAAGCAAGCAAGCTTCAGGC 7235
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      100 Y.....SerArgSerTyrGlnValTyrCysPheAsnLeuLysLysAlap 115
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7236 TATCTCAACAGCAATGAGGAGGTTTCAC.....AAGAGGCA. 7272
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      115 heProLeuGlySerAspSerValLysLysTyrTyrLysLysHisasp 131
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7273 .....AACTGTGTGTGTCAGTGGCTGGAATCAAAAGAG 7305
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      132 GlyLleSerThrLysPheGluAspTyrAlaMetSerProAlaGleThrG1 148
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7306 GAATGC...CTGAATTCATGTATGCCATCTCATCTCCACACAGACAGA 7352
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      148 YAsp.....GluLeuAsnGlnLysLeuArgAlaValMet 160
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7353 AACAGTGAAGCCCAAGCTGAATCTAACAAAGCCCTCTCGCTGAGTTG. 7401
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      160 YrAsnGlyHisProGlnAsnAlaAsnGlyLleMetGluGlyLeuGluPro 176
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7402 .....GAACAGAAATTCTCCAAAATTCAAAAGTAAGAAAGAACCC 7440
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      177 LeuAsnAlaIleArgValThrGlnGluAlaValTrpTyrTyrSerAspAs 193
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      7441 CTGGCTGTGATTACTGCTGACATAT..... 7464
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      193 nAlaProIleSerAsnProAspGluSerPheLysArg..... 205
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7465 ....CCCAACTCACAGGAAAGCAAAATTTGGAAGAAATTCAGAGAAGAAC 7510
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      205 ..... 205
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7511 TCAATTCCTCCGATGGAAAGGCCACTGAGTTACTGTGGCTGGCAAGG 7560
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      206 .....GluSerGluSerAsnLeu.....ValSerThrSe 215
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      7561 CAGTAGAGGAAATCTGCAATCATCTGCGCTGCCCTCCAGCTGCAGATC 7610
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      215 rGlnLeu.....SerLeuMetArgGlnAlaLeuL 225
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      7611 CCAGCTCCAGCCGCTGGCTGATGAGAAAGAACTATGATGGAGTGGCTGG 7660
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      225 YsGlnLeu...IleAspProAsnLeuAlaThrLysMetProLysGlnVal 240
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      7661 GGGCCCTGTCTATTGACCCCAACATGTTGAATGCACAAAGCAACAGGCTC 7710
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      241 ProAspAspPheGlnLeuSerIlePheGluSer...GluAspLysGlyAs 256
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7711 .....CAGTTTATGCTTAAGAAATTTGAACACAGCAAGCAACGATGA 7754
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      256 pLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyGlyLeuValProT 273
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7755 GCAACTGAATGAGCAGCTCAGGCAATCTTAACAGGC..... 7791
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      273 hrLysProProThrProGlyAspProPheMetProProAsnGlnProGln 289
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      7792 .....CTTGAGATGTCTCTGTCTCACCACGCAAGTACAG 7837
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      290 ThrThrSerValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLe 306
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      7828 AAAGAACTCCGAGACATCAATCAG..... 7851
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      306 uLeuGluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheG 323
          : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

      7852 .....AAATGGGTTGAGCTGACT...GACAACTCAACTCCCGTT 7888
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      323 LnaIaArgValPheSerSerAsnAspIleGlyGluArgIleGluLeuSer 339
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      7889 CCAGCCAAATTT.....GACCACCTATTGTTAAGACACCCCACTACCAG 7932
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      340 AspGlyThrTyrThrLeuThrGluLeuAsnSerProAlaGly.....Ty 354
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      7933 GAATGCTCCAGCACTTATATCAGACAGAGGTGAGGAGCTGTGACAAAGGCT 7982
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      354 rSerIleAlaGluProIleThrPheLysValGluAlaGlyLysValTyrT 371
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      7983 GAGGTCCAGTCAGCTATCAGCACCCCAACCAAGAGGCTGTAAAG..... 8025
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      371 hrIleIleAspGlyLysGlnIleGlnAsnProAsnLys..... 383
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8026 .....CAGCAATGTGAAGAGACCACTGAATTCGATCTGCAC 8061
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      384 ..GluIleValGluProTyrSerValGluAlaTyrAsnAspPheGluG1 399
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      8062 TTGGAGCAGTTAGACCAAGAGGTTAAGAGGCTCAGACACTGTGCATGA 8111
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      399 uPheSerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAlaLysA 416
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8112 ACTCTAGTGTCTCATTTGTGTGAGCAGTACCTCAAGATGAACGTGAAGAGC 8161
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      416 sn.....LysAsnGlySerSer 421
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8162 GTTTGGAGACAGTGGCCCTGCTCCAGAGTTTGAAGACCTTGCAAGCC 8211
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      422 GlnValValTyrCysPheAsnAlaAspLeuLysSerProAspSerG1 438
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8212 GATCGCAATTAAACAGACTCCAGCAGCTCTTCGACAGCCAGAGTTCCA 8261
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      438 u.....AspGlyGlyThrMetThrProAspPheThrGlyG 452
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8262 GCAATGTTTGTATGATGTGAGGACCTGTTGATGAATAAACAAGCCAGC 8311
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      452 LuValLysTyrThrHisIleAlaGlyArgAspLeuPheLysTyrThrVal 468
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8312 AAGCAAAAACCTGCCCAATTTCTGCAAA..... 8340
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      469 LysProArgAspThrAspProAspThrPheLeuLysHisIleLysVal 485
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8341 .....TTGAGCGGCTACAGTCTCA 8360
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      485 ILeGluLysGlyTyrArgGlyLysGlyGlnAlaIleGluTyrSerGlyL 502
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8361 GCTACAGAGAGAAAGAGATTTCAGAAAAGCTTAATCAACACAGTGGCT 8410
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      502 euThrGluThrGlnLeuArgAlaAlaThrGlnLeuAlaIle.TyrTyrPh 518
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      8411 CCTATGAGGTATTTGTGCTGAAGGGAATCTCTACTTCTTCTGTAAGCT 8460
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      518 euThrAspSerAlaGluLeuAspLysAspLysLeuLysAspTyrHisGlyP 535
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8461 CCTG.....GAGAAGAGAAAGAAAGACTCTACAAAA 8489
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      535 heGlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGlu 551
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8490 CCAGTTGTTGAGCTCAAAAAC.....ATTGGAAAGAGCTTGTAATA 8533
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      552 TyrAlaGlnAspSerAsnProProGlnLeuThrAspLeuAspPheIle 568
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      8534 AAACGTGCAAGACAAATCCAGC.....CTCAAGATTTTAT 8570
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      568 eProAsnAsnAsnLysTyrGlnSerLeuIleGlyThrGlnThrPheProG 585
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      8571 GCAGAAAGCTCAGAAATAT.....CAGTGGCATGTGG 8602
          : : : : : : : : : : : : : : : : : : : : : : : : : :
      585 LuAspLeuValAspIleIle.....ArgMetGluAspLys 596
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      8603 AAGACCTTGTGCTCATGATAGAAAGATTGTAAGCTAAGATGTGAGATTG 8652
          : : : : : : : : : : : : : : : : : : : : : : : : : :

```





## TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 610-5790

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:

LENGTH: 13794 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 54:

US-08-956-171C-54

## alignment\_scores:

Quality:	123.50	Length:	668
Ratio:	0.468	Gaps:	29
Percent Similarity:	39.521	Percent Identity:	20.958

## alignment\_block:

US-09-494-297-2 x US-08-956-171C-54/rev ..

Align seg 1/1 to reverse of: US-08-956-171C-54 from: 1 to: 13794

```

153 GlnLysLeuArgAlaValMetTyrAsnGlyHisProGlnAsnAlaAsnG1 169
      |||||
8153 CAAAGACAACTACAGTAGAAGAAATGGAAATCAGCTACTGATATTA 8104
169 YLeMetGlu.....GlyLeuGluProLeuAsnAlaLeaArgValT 183
      |||
8103 AACAGGTAAACACAAACAACTGCACTAAGCTTAATCATATAGAGAAA 8054
183 hrgGlnAlaValITrPTyTrSerAspAsnAlaProIleSerAsnPro 199
      |||||
8053 CTCAA.....TCATATACGCAACAGTACA..... 8028
200 AspLysSerPheLysArgGluSerGluSerAsnLeuValSerThrSerG1 216
      ::::: |||::: |||::: |||::: |||::: |||:::
8027 .....GAACAACCTCAAACGCCACACACAGTAACACTGAGA 7990
216 nLeuSerLeuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeuA 233
      : |||||
7989 AGCACCAAAACAGTACAGCACCAAACTGCACAAACAGCAATATAG 7940
233 IaThrLysMetProLysGlnValProAspAspPheGlnLeuSerIlePhe 249
      |||||
7939 AACACGTTAAAGAGAGTACTTAAGAGAAAGCAAACTCAAGTTAAG 7890
250 GluSerGluAspLysGlyAspLysTyrAsnLysGlyTyrGlnAsnLeuLe 266
      |||||
7889 GAACACACACATCTCAAGC.....AATAGCGAGATCAAAGCAAGT 7846
266 uSerGlyLeuValProThrLysProProThrProGlyAspProProm 283
      ::::: |||||
7845 AGAT.....TTAACCTTAATAAG.....G 7826
283 etProProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAla 299
      |||||
7825 CTACACAAAATCAAGTCGCACAAACAGTTGAAGTGCACACGCCAAGA 7776
300 IleGlyAspTyrSerLysLeuLeuGluGlyAlaThrLeuGlnLeuThrG1 316
      ::::: |||||
7775 ACGGCAATCAAGAACTAAG..... 7758
316 yAspAsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleG 333
      |||||
7757 .....CCACGTGACAAAGATCAGCAGATGTAG 7730
333 LysLysArgIleGlnLeuSerAspGlyThrTyrThrLeuThrGlnLeuAsn 349
      ::||| |||
7729 CGGAAGCTAAGGAAGCTTAGAACGG..... 7704
350 SerProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAl 366
      |||||

```

```

7703 .....AAAGTGAAC 7693
366 aGlyLysValTyrThrIleLeuAspGlyLysGlnIleGluAsnProAsnL 383
      :|||
7692 GGGTACAGATGTACA.....ACTA 7673
383 ysgLileValGluProTyrSerValGluAlaTyrAsnAspPheGluGlu 399
      |||||
7672 AAGTTACAGTAGAAATGGTTCTATGAGGGCCATACAAATACAAATAA 7623
400 PheSerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAlaLysAs 416
      ::::: |||||
7622 GTAGAACCTCATGCAGACACAGACGGCTACTAAATATAGTTGAATTT 7573
416 nLysAsnGlySerSerGlnValValTyrCysPheAsnAlaAspLeu... 431
      ::||| |||
7572 TGAGAAATGTTACATCAAGGTGACTAC...TTTGACTTCTTTATACAA 7526
432 .....LysSerProProAspSer 437
7525 ATTAATGTAAATACGCAATGCGCTATCACTGCTAGAAAAGTCCAGAAATT 7476
438 GluAspGlyGlyLysThrMetThrProAspPheThrGlyGluValLys 454
      ::::: |||||
7475 AAAAATGGTTACGTCGTAATG.....GCGACAGGTGAAGTTT 7438
454 sTyThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysProA 471
      ::|||
7437 AGAA.....GGTGAAGATTAGATATACATTTCAATGATATTGTG 7397
471 rGAspThrAspProAspThrPheLeuLysHisIleLysLysValIleGlu 487
      ::|||
7396 AAGATAAAGCTTGATGATACGGCTGAAGTGAATTAATTAATTTATTTGAT 7347
488 ..LysGlyTyrArgGluLysGlyGlnAlaIleGluTyrSerGlyLeuTh 503
      |||
7346 CCTAAACGTACAAACATATGAATGAATCAAACTATACACTCACTAA 7297
503 rGluThrGlnLeuArgAlaAlaThrGlnLeuAlaIleTyrTyrPheThrA 520
      :||| |||
7296 TGAAGACAA.....A 7286
520 sPseRAlaGluLeuAsp.....LysAspLysLeuLysAspTyrHis 533
      |||||
7285 CTTCAAAGGAATTAGATGTTAAATATTAAGATGATTTGGAAATATAT 7236
534 Gly.....PheGlyAspMetAsnAsp.... 540
      ::::: |||||
7235 GCCAATTTAATGATGATGATGACATTTAATAAGCAATATATGATT 7186
541 .....SerThrL 543
7185 TTCGCAATGTCATTATTAAACCTAATATAGTAAACGACAAAGTGTGA 7136
543 euAlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProPro 559
      ::|||
7135 CTGTTACTGGAACCTTAATGAAGGTAGTAATTCAGAAATGAATCAACA 7086
560 GlnLeu..... 561
7085 AAAGTTAGCATATTTGATACTTGCGTAAATAATGAAGCATACGAAAGAG 7036
562 .....ThrAspLeuAspPhePhe..... 567
7035 TGTATATGCAAAATACGACGATACTCTTAATTAAGACGTACACAACTA 6986
568 .....IleProAsnAsnAsnLysTyr..... 574
6985 ATATAGATGGGAATTTGATTTACAAATTAATGGAAGCTATTCATTGAT 6936
575 .....GlnSerLe 577
6935 ATAGAAATCTAGATAAATACTTATGTTGTTCACTATGATGAGAGATATT 6886

```

```
577 uilegIYThrGln.....TrpHisProG 585
| |||||:::
6885 AATGGTACGTAGTATGATTGTAACACAAAGCTAGACATCCAG 6836
|||
585 lnspleu.ValaspIleIleargMetGluaspLysGluValIlePr 601
|||:::|||||:::|||||
6835 ACCAAGCTTATTAGTATTATTATGATAGAGAT.....ATACC 6798
|
601 O.....ValThrHisAsnLeuThrL 608
| |||:::
6797 TTACTTGGATATAGTGTATTATACGTAATAAAGCAAGGAA 6748
|
608 euArgLysThrValThrGlyLeuAlaGlyAspArgThrLysAspPheHis 624
:::|||||
6747 TGAGAAAAATGTCGCAT..... 6729
|
625 PheGluIleGluLeuLysAsnAsnLysGlnGluLeuLeuSergInThrVa 641
:::|||||
6728 .....ATTCAAATATATAATTGAAATATTAAGAAAGATACAT 6691
|
641 lLysThrAspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleA 658
|||:::|||||:::|||||
6690 TAAAGAAACTCTTACAGCTCAATATGATAGAAATTAGTAACTACTGTG 6641
|
658 snLeuLysHisGlyGluSergLeuThrLeuGln..... 668
:::|||||
6640 AAGAGCAATAT...GATTCATCAACTCTTGACATGATACACACACGCT 6594
|
669 .....GlyLeuProGluGlyTyrSeryLeuValLysG 680
|||:::|||||
6593 ATAGATGCTGAGGTGATATGTTGATGATACATGAAACAATAGAGA 6544
|
680 uThrAspSergLugly.....TyrLysValLysValAsnSerg 693
|||||:::|||||
6543 AAGGATTCATCAGCTATGTGATATGATCCATTCGTGTGGATAGCG 6494
|
693 ln 693
::
6493 AA 6492
```